STRUCTURAL, DYNAMIC AND INTERACTION STUDIES OF AUTOPHAGY-RELATED PROTEINS

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Biochemistry

in the School of Biological Sciences

by Arvind Chand Ravichandran

University of Canterbury

March 2016

ABSTRACT

Autophagy is a crucial cellular catabolic process that degrades unwanted cytoplasmic material (cargo) by engulfing the material into a double-membrane vesicle (autophagosome) and then presenting it to the lysosome. The work in this thesis explores interactions between the proteins involved in the autophagy pathway.

Focal adhesion complex Interacting Protein of 200 kDa (FIP200) and autophagy related protein 13 (Atg13) are members of the autophagy initiation complex. Attempts were made to locate the region on FIP200 that binds the Atg13 by dividing FIP200 into six fragments of equal length and performing pull-down assays with Atg13. However, because of a false positive interaction observed for all six FIP200 fragments, binding between the Atg13-interaction region on FIP200 and Atg13 could not be identified.

Pleckstrin homology domain-containing protein family member 1 (PLEKHM1) is a selective autophagy receptor that interacts with the autophagy-related protein 8 homologues (Atg8 proteins: LC3A, LC3B, LC3C and GABARAP, GABARAPL1, GABARAPL2). The affinity of PLEKHM1 for the LC3 family is higher than that for the GABARAP family of proteins. The Atg8 proteins are involved in tethering cargo to the autophagosome. The crystal structures of five human ATG8 proteins were solved in complex with a peptide with the sequence of the PLEKHM1 LC3-Interacting Region (LIR) to understand the structural basis of their differing binding affinities. These structures demonstrated that residues Arg28 and Glu17 of the GABARAP family proteins interact with Asn637 and Trp635 residues of the PLEKHM1 LIR peptide. However, these interactions were either weak or not present in the LC3 family proteins, suggesting that they may contribute to the higher binding affinity for the PLEKHM1 LIR observed with GABARAP proteins. In addition, a structural comparison of the human ATG8 proteins suggested that the “W-site” hydrophobic pocket, where Trp635 in PLEKHM1 binds, is deeper in GABARAP family proteins than in LC3 family proteins. Together, the structural analysis provided a rationale for varying binding affinities between GABARAP and LC3 family proteins.

The serine/threonine-protein kinase ULK1, which is part of the autophagy initiation complex, also interacts with the Atg8 family proteins. Using surface plasmon resonance, molecular dynamic (MD) simulations, mutational analysis and X-ray crystallography, the interaction of the ULK1 LIR with GABARAP and LC3A was studied to understand the difference in the LIR interaction mechanism. Structural comparison of the LC3A and GABARAP proteins in their LIR-bound and unbound states showed that LC3A shows more variation in the structure of its "W-site" hydrophobic pocket between its bound and unbound states than the "W-site" of GABARAP. This was verified by MD simulations. Additionally, the MD simulations suggested an electrostatic interaction between LC3A Arg70 and ULK1 Met359. This was demonstrated by the decrease in binding affinity when Arg70 was mutated to leucine. In contrast, mutation of Arg67 (which is equivalent to Arg70 in LC3A) to leucine did not have any effect on LIR binding.

Overall, this work provides key insights into the differences in the mechanisms of interaction demonstrated by the GABARAP and LC3 family proteins, which is an initial step towards understanding the different functions of the members of the two human ATG8 subfamilies (GABARAP and LC3).

ACKNOWLEDGEMENTS

First and foremost, I thank my supervisor Dr. Renwick Dobson for giving me this opportunity. I couldn't have asked for a more supportive and understanding supervisor. Thank you, Ren, for your unfailing guidance, encouragement and not giving up on me during my difficult times, and for all the Skype meetings we had during my stay in India. I would also like to thank Dr. Hironori Suzuki for his guidance, teaching, support and the amazing Japanese lunch sessions, and Prof. Conan Fee for kindly letting me use the surface plasmon resonance equipment, and for his useful insights. I would like to acknowledge Dr. Grant Pearce and my collaborators from Prof. Ivan Dikic's lab for their valuable input and support. This work would not have been possible without our amazing lab managers, Jackie Healy and Rayleen Fredericks.

During these four years, my colleagues were my friends. I thank each and every member from the Dobson lab and all my colleagues/friends from BIC 6th floor: Kat, Rachy, Jen, Chris, Muge, Amy, Eric, Jeremy, Moritz, Sarah, and Ali. I couldn't have done this without your support. Staying away from home for more than three years was tough, but my Indian friends in New Zealand made sure I never missed home. Thank you Nivas, Mangai, Akshita, Rishi, Amol, Achinta, Nilanjana, Shaurabh, Ishan, Arun and Vivek.

Finally, I thank my amazing family, friends and relatives in India for their unconditional support. My parents, for their love, encouragement and belief in my abilities. I could not have completed this journey without their blessing and prayers. I thank my dad for his constant support, guidance and enthusiasm, and my mum for the emotional support and care she has provided.

ABBREVIATIONS

A_{280}	absorbance at 280 nm
ALFY	autophagy-linked FYVE protein
ATG	autophagy-related genes
Atg101	autophagy-related protein 101
Atg12	autophagy-related protein 12
Atg13	autophagy-related protein 13
Atg5	autophagy-related protein 5
Atg7	autophagy-related protein 7
Atg8	autophagy-related protein 8
AU	absorbance unit
B-factor	temperature factor
BECN1	beclin 1 autophagy-related protein gene
CMA	chaperone-mediated autophagy
CV	column volume
Cvt	cytoplasm to vacuole targeting pathway
DNase	deoxyribonuclease
dNTPs	deoxyribonucleotide triphosphates
DSF	differential scanning fluorimetry
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ε_{280}	extinction coefficient
FIP200	focal adhesion complex interacting protein of 200 kDa
GABARAP	gamma-aminobutyric acid type A receptor-associated protein
GABARAPL1	gamma-aminobutyric acid type A receptor-associated protein-like 1
GABARAPL2	gamma-aminobutyric acid type A receptor-associated protein-like 2
GFP	green fluorescent protein
GST	glutathione S-transferase
HEK cells	human embryonic kidney cells
HEPES	N-2-hydroxyethlpiperazine-N'-2-ethane sulphonic acid
IPTG	isopropyl β-D-1-thiogalactlopyranoside
*K*d	dissociation constant
LB	Luria Bertani
LDS	lithium dodecyl sulphate
MAP1LC3A	microtubule-associated proteins 1A/1B light chain 3A
MAP1LC3B	microtubule-associated proteins 1A/1B light chain 3B
MAP1LC3C	microtubule-associated proteins 1A/1B light chain 3C
MD	molecular dynamics simulation
MES	2-(*N*-morpholino)ethanesulfonic acid

MOPS	3-morpholinopropanesulfonic acid
mTOR	mammalian target of rapamycin
MW	molecular weight
MWCO	molecular weight cut-off
NBR1	neighbour of BRCA1 gene 1
NDP52	nuclear domain 10 protein 52
NHS	*N*-hydroxysuccinamide
NTA	nitrilotriacetic acid
OD_{600}	optical density at 600 nm
OPTN	optineurin protein gene
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PBST	phosphate buffered saline with Tween-20
PCR	polymerase chain reaction
PDB	protein data bank
PE	phosphatidylethanolamine
PEG	polyethylene glycol
PI3K	phosphatidylinositol 3-kinase
PLEKHM1	Pleckstrin homology domain-containing protein family member 1
R_{factor}	residual factor
R_{free}	free R_{factor}
RFU	relative fluorescence units
RMSD	root mean square deviation
RMSF	root mean square fluctuation
rpm	revolutions per minute
RU	response unit
S200	Superdex 200 matrix
SDS	sodium dodecyl sulphate
SEC	size exclusion chromatography
SLS	static light scattering
SOC	super optimal broth with glucose added
SPR	surface plasmon resonance
Tris	tris(hydroxymethyl)aminoethane
U	unit
ULK1	unc-51-like autophagy activating kinase 1
UPP	ubiquitin proteasome pathway
UV	ultraviolet
v/v	unit volume per unit volume
V_o	void volume
w/v	unit weight per unit volume
w/w	unit weight per unit weight

Table of contents

Chapter 1. INTRODUCTION

1.1 Background

Cells are the most basic unit of life. They transcribe and translate the genetic code into proteins, the workhorses of life. Apart from providing structure to the cells, proteins perform various important functions, such as catalysing biochemical reactions, cellular signalling and communication, transporting molecules, defending against foreign invaders, and providing contraction and movement. It is essential that the right quantities of proteins are maintained in a properly functional and folded state.

Autophagy (from the Greek words *auto*, meaning "self", and *phagein*, meaning "to eat") is a eukaryotic cellular pathway that degrades and recycles unwanted cytoplasmic materials into simple amino acids. Autophagy plays a critical role in striking a balance between protein synthesis and protein degradation (Mitch & Goldberg, 1996). When survival becomes strenuous, unwanted and long-lived proteins and organelles are degraded by autophagic pathways into simple amino acids, which are then recycled for synthesis of new proteins to maintain cellular homeostasis.

Autophagy is a carefully coordinated pathway involving many protein partners and locations within the eukaryotic cell. The following literature review begins with an overview of the various pathways involved in protein degradation, followed by a detailed description of autophagy.

1.1.1 Types of protein degradation

Proteolytic activity in cells was first detected in 1898 (Hahn, 1898). Following the discovery of the lysosome, researchers focused on the field of protein degradation (De Duve, 1963). In eukaryotic cells, protein degradation can be broadly classified into two separate pathways (Figure 1.1): the non-lysosomal pathway (i.e. the ubiquitin-proteasome pathway, which is briefly reviewed here), and the lysosomal protein degradation pathway (which includes autophagy and will be reviewed in-depth).

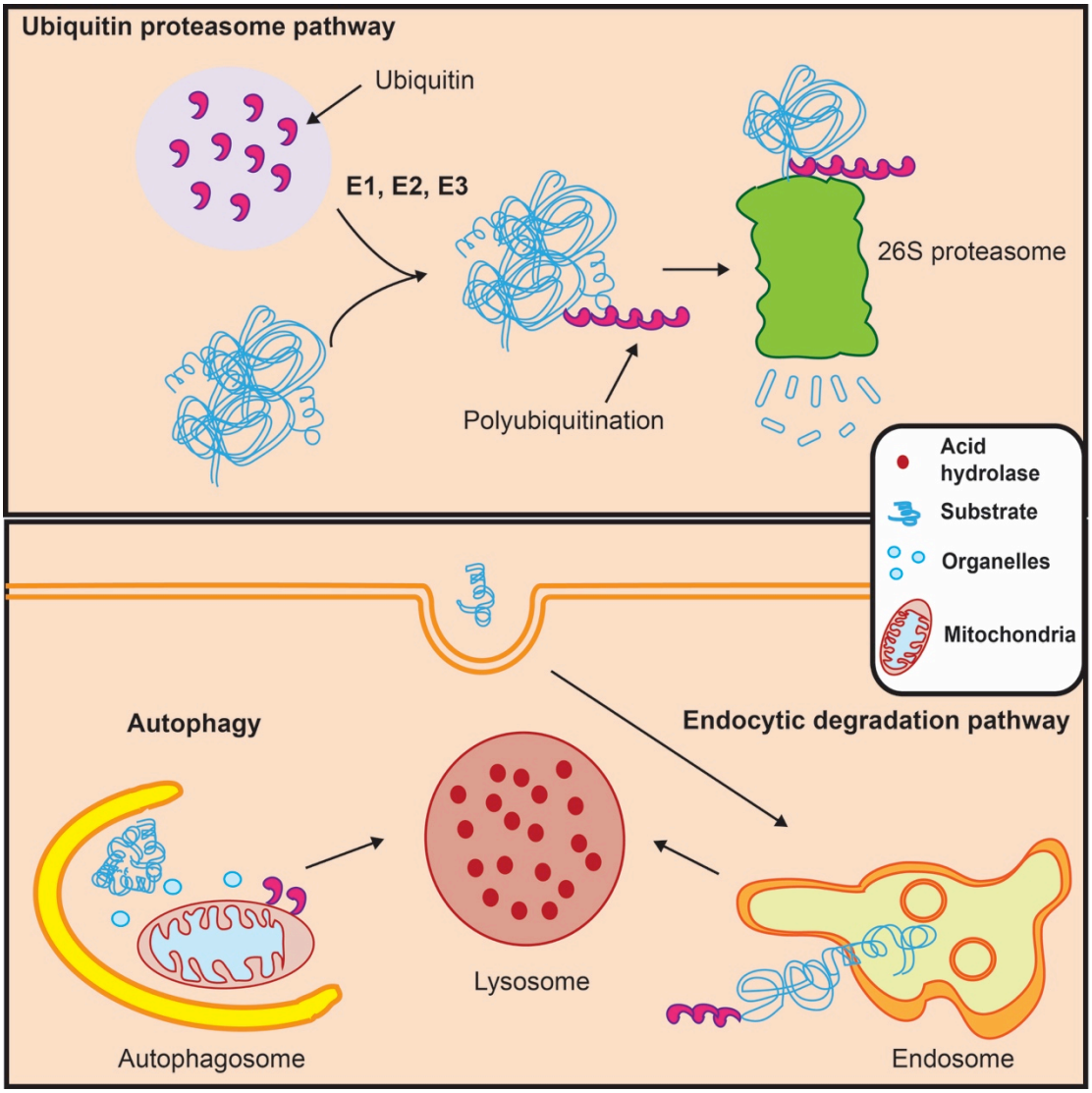


Figure 1.1: The two protein degradation pathways in eukaryotic cells: The ubiquitin proteasome pathway degrades ubiquitinated protein via the proteasome. The lysosomal pathway mainly degrades protein through endocytosis and autophagy.

1.1.1.1 Non-lysosomal ubiquitin-proteasome pathway

The ubiquitin proteasome pathway (UPP) degrades short-lived and aberrant intracellular proteins by covalently attaching a small 76-amino acid protein, called ubiquitin, to the protein that is to be degraded (Heinemeyer *et al.*, 1991) (Figure 1.1). Tagged proteins are recognised by the 26S proteasome, a large multi-catalytic protein complex that degrades ubiquitinated proteins into small peptides (Baumeister *et al.*, 1998). Ubiquitin is attached to the targeted protein via three enzymatic reactions (mediated by the enzymes E1, E2 and E3), in which E1 activates ubiquitin, E2 carries ubiquitin and prepares it for conjugation and E3 identifies the protein to be degraded and ligates it to the activated ubiquitin (Lecker *et al.*, 2006). Often, a protein is tagged with a chain of more than one ubiquitin in a process called poly-ubiquitination.

The 26S proteasome is comprised of approximately 60 subunits. It has two main components: the 20S core particle, which is the central hollow cylinder where the ubiquitin-labelled protein is degraded, and the 19S regulatory particle, located on either side of the cylinder, where the protein to be degraded is unfolded and the ubiquitin tag is detached for reuse (Figure 1.1). Dysfunction in the UPP can lead to cancer, neurodegenerative disorder, immune disorders and chronic kidney disease (Cooper, 2000a, Hochstrasser, 2009, Lecker *et al.*, 2006).

1.1.1.2 Lysosomal protein degradation pathway

Cells contain closed membrane vesicles, called lysosomes, that store approximately 50 different acidic hydrolases involved in the degradation of a wide range of materials, including proteins, nucleic acids, organelles, lipids and carbohydrates. These enzymes are active inside the lysosome at pH ~5.

One of the main functions of lysosomes is to degrade extracellular proteins that are transported into the cell through endocytosis. In addition to the endocytic

pathway, lysosomes also degrade materials via two other pathways: phagocytosis and autophagy. In phagocytosis, the phagocytic vesicles (called phagosomes), engulf large particles, such as bacteria and cell debris, and then fuse with the lysosome to form the phagolysosome. The phagolysosome degrades the materials that have been taken up into the vesicle (Cooper, 2000b). The main focus of this thesis will be autophagy, as explained below.

1.1.1.3 Autophagy

Autophagy is a bulk degradation process that is conserved in eukaryotic cells (De Duve & Wattiaux, 1966). Under cellular stress conditions, such as starvation, autophagy is used to degrade misfolded proteins, long-lived proteins and damaged organelles by sequestering them in a double-membraned vesicle, and then fusing this vesicle with the lysosome for degradation (Mortimore & Schworer, 1977). Dysfunction in autophagy leads to certain types of cancer (colorectal cancer, prostate cancer and lymphoma) (Liang *et al.*, 1999, Qu *et al.*, 2003, Yue *et al.*, 2003, Mariño *et al.*, 2007, Takamura *et al.*, 2011, Morselli *et al.*, 2011), neurodegenerative disorders (Berger *et al.*, 2006, Ravikumar *et al.*, 2002, Webb *et al.*, 2003, Shibata *et al.*, 2006, Menzies *et al.*, 2010, Williams *et al.*, 2008, Zheng *et al.*, 2010), immunological disorders (Lee, Mattei*, et al.*, 2010, English *et al.*, 2009, Li *et al.*, 2008, Campbell & Spector, 2011, Schmid *et al.*, 2007), cellular ageing and cell death (Lipinski *et al.*, 2010, De Kreutzenberg *et al.*, 2010, Caramés *et al.*, 2010, Lee *et al.*, 2009). Autophagy can be categorised into microautophagy, chaperone-mediated autophagy (CMA), and macroautophagy.

1.1.1.3.1 Microautophagy

Microautophagy, which is best characterised in yeast, is a process where the cytoplasmic material is engulfed by direct invagination into the lysosomal membrane and then degraded (Ahlberg & Glaumann, 1985). Two separate multistep pathways have been identified in microautophagy: the type I pathway, in which soluble components are degraded, and the type II pathway, in which selected organelles are degraded (Kunzt *et al.*, 2004). In both pathways, budding is initiated by the formation of a structure called the "autophagic tube" (Kunzt *et al.*, 2004). This is followed by vesicle formation, expansion, scission, degradation and recycling (Figure 1.2). Improper functioning of microautophagy can lead to various human diseases, such as Creutzfeldt-Jakob disease, Gerstmann-Straussler-Scheinker, Huntington's disease, albinism and Pompe disease (Kunzt *et al.*, 2004, Li *et al.*, 2012).

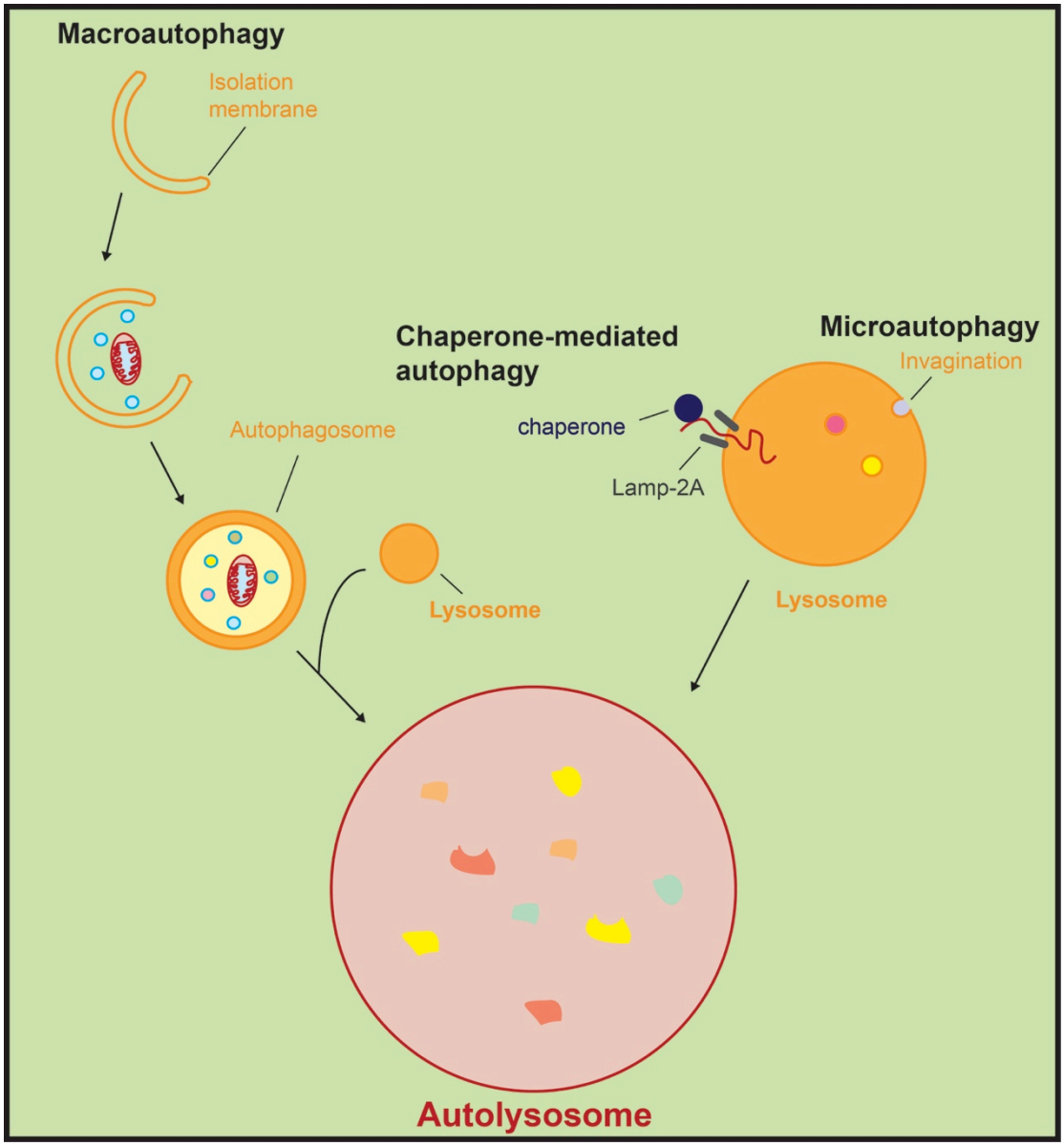


Figure 1.2: Overview of the three types of autophagy. Macroautophagy involves sequestration of cytoplasmic material in a vesicle called an autophagosome, and fusion with a lysosome. In microautophagy, cytoplasmic material buds into a lysosome to form a vesicle before being degraded. Finally, in chaperone-mediated autophagy, chaperones present the protein directly to the lysosome to be degraded.

1.1.1.3.2 Chaperone-mediated autophagy (CMA)

CMA is best understood in higher eukaryotes, such as mice and humans (Kaushik & Cuervo, 2012). Proteins that are to be degraded are recognised by the Hsc70 chaperone and presented to the lysosome for degradation. The chaperone recognises a specific peptide motif (KFERQ or a KFERQ-like sequence) and binds to

the protein. The protein is delivered to the lysosome by lysosome-associated membrane protein type 2A (LAMP-2A). Unlike other lysosomal pathways, CMA does not require vesicle formation; instead, it presents the cargo directly to the lysosome (Figure 1.2). Malfunction in the CMA pathway leads to neurodegenerative disorders such as Parkinson's disease (Cuervo & Wong, 2014), kidney disease (Cuervo, 2010) and lysosomal storage disorder (Cuervo & Dice, 1996).

1.1.1.3.3 Macroautophagy

Macroautophagy, the pathway at the centre of this work, is the best-understood autophagic pathway. Autophagic vesicles (autophagosomes) were first observed by Christian deDuve in 1966 (De Duve & Wattiaux, 1966).

Macroautophagy is a stepwise process. It can be categorised into the following stages: initiation, elongation, maturation and fusion. During the initiation stage, a double membrane, called the phagophore, is formed from the isolation membrane (Figure 1.2). The isolation membrane is a cup-like, double-membrane structure floating in the cytoplasm. The source of the isolation membrane is poorly understood. Various studies suggest that the double membrane is obtained from the smooth endoplasmic reticulum, the plasma membrane or the mitochondria (Hailey *et al.*, 2010, Ravikumar *et al.*, 2010, Hayashi-Nishino *et al.*, 2009). The elongation phase begins with the addition of the double membrane to the phagophore and sequestration of unwanted cytoplasmic material. Phagophore elongation leads to closure and formation of a vesicle called an autophagosome. The autophagosome is transported to the lysosome during the maturation phase, then the two bodies fuse to form an autolysosome in the fusion phase. Finally, the autolysosome degrades the sequestered materials inside the vesicle (Metcalf *et al.*, 2012). Interest in macroautophagy

increased after the genes involved in this pathway were identified. Autophagy is regulated by the autophagy-related (ATG) genes, first reported by Ohsumi's group in 1997 (Matsuura *et al.*, 1997). Soon after, Mizushima identified mammalian counterparts of the yeast proteins Atg5 and Atg12 in mammals (Mizushima *et al.*, 1998).

1.2 Autophagy in disease

Dysfunction of autophagy leads to various disorders, making it an interesting and compelling target of study. Most neurodegenerative disorders are caused by the accumulation of protein aggregates or toxic proteins in neurons (Rubinsztein, 2006). Studies in mice, *Drosophila* and zebrafish have shown that upregulation of autophagy results in reduced levels of these protein aggregates and toxic proteins, making the pathway an attractive drug target (Shibata *et al.*, 2006). In other studies, upregulating autophagy reduced the levels of proteins encoded by genes involved in disorders such as obesity (Codogno & Meijer, 2010, He *et al.*, 2012), cancer (Bursch *et al.*, 1996) and encephalitis (Liang *et al.*, 1998). For certain types of cancer, including colorectal cancer, prostate cancer and lymphoma, inhibiting autophagy is being considered as a therapeutic strategy (Janku *et al.*, 2011, Yang *et al.*, 2011). Induction of autophagy also increases longevity, making it a possible target for ageing and longevity research (Rubinsztein *et al.*, 2011). The roles of autophagy in various disorders are discussed in detail below.

1.2.1 Autophagy and metabolic disorders

Lipophagy, an alternative lipid metabolism pathway, uses lysosomes to degrade liver lipid droplets (Liu & Czaja, 2013). The free fatty acids produced from

this pathway are oxidised by mitochondria (Singh, Kaushik*, et al.*, 2009). The ATG7 gene encodes E1-like ligase, which is involved in the formation of the autophagosome. ATG7 knock-out mice show an increase in hepatic lipid levels that leads to obesity (Yang *et al.*, 2010). Restoration of this gene decreased obesity and insulin resistance, and improved glucose tolerance (Zhang *et al.*, 2009, Singh, Xiang*, et al.*, 2009). Furthermore, a decrease in lipophagy was observed in both genetic and diet-induced obese mouse models (Yang *et al.*, 2010, Zhang *et al.*, 2009, Singh, Xiang*, et al.*, 2009). Mouse models with increased autophagy in response to exercise are protected against impaired glucose tolerance and triglyceride levels induced by a high fat diet (He *et al.*, 2012). Such beneficial effects were not noticed in mouse models where autophagy was not induced in response to exercise (Rubinsztein *et al.*, 2012). These findings suggest that autophagy plays a beneficial role in obese patients. Thus, the autophagy pathway is an attractive drug target to treat various metabolic disorders like obesity and diabetes (Codogno & Meijer, 2010).

1.2.2 Autophagy and neurodegenerative disorders

Accumulation of protein aggregates in neurons leads to various neurodegenerative disorders, such as Parkinson's disease (Berger *et al.*, 2006), Alzheimer's disease (Lee, Yu*, et al.*, 2010, Pickford *et al.*, 2008) and Huntington's disease (Shibata *et al.*, 2006). Most of these protein aggregates are autophagy substrates. Clearance of the toxic mutant protein huntingtin, which causes Huntington's disease, is slowed down when autophagy is impaired (Shibata *et al.*, 2006). In contrast, upregulation of autophagy enhanced clearance of this toxic protein (Shibata *et al.*, 2006). Similar studies performed on other mutant proteins, such as mutant α-synuclein, which causes Parkinson's disease, and mutant ataxin 3, which

causes spinocerebellar ataxia type 3, showed that upregulation of autophagy had beneficial effects (Berger *et al.*, 2006, Ravikumar *et al.*, 2002, Webb *et al.*, 2003).

Because defective autophagosome formation can cause neurodegenerative disorders, drugs that upregulate autophagosome formation can be used to treat these diseases. For example, Lafora epilepsy is caused by mutant laforin, which activates the mammalian target of rapamycin (mTOR) complex (see Section 1.4.1) (Aguado *et al.*, 2010). This complex, in turn, negatively regulates autophagy by impairing autophagosome formation (Hosokawa *et al.*, 2009). Therefore, drugs like rapamycin can be used to inhibit mTOR and activate autophagy to cure this disease (Aguado *et al.*, 2010). Evaluating the proteins involved in autophagy for their potential as drug targets may allow researchers to develop cures for these disorders.

1.2.3 Autophagy and cancer

Interestingly, autophagy can be a tumour suppressor, as well as help in the survival of tumour cells during metabolic stress and chemotherapy. For example, ATG gene *BECN1*, which encodes beclin 1, is part of the phosphatidylinositol-3-kinase (PI3K) complex that mediates vesicle-trafficking processes in autophagy (Su *et al.*, 2015). Deletion of *BECN1* increases susceptibility to breast, ovarian and prostate cancer in humans (Liang *et al.*, 1999, Yang *et al.*, 2011). It has also been shown that tumour suppressor genes activate autophagy, which might in turn inhibit tumour initiation. For example, tuberous sclerosis 1 (TSC1) and tuberous sclerosis 2 (TSC2) inhibit mTOR, which in turn activates autophagy. On the other hand, tumour signalling pathways such as oncogenic receptor tyrosine kinase, class I PI3Ks and protein kinase B (AKT) inhibit autophagy by activating mTOR (He & Levine, 2010, Pimkina *et al.*, 2009). Thus, autophagy represents a useful drug target for cancer.

Further research and clinical trials will help to shift the focus on autophagy for cancer therapy (Rubinsztein *et al.*, 2012).

1.2.4 Autophagy and infectious disease

The first evidence of the involvement of autophagy in infectious disease was reported in 1998, when Liang *et al*. showed that expression of the ATG gene *BECN1* protected mice against alphavirus encephalitis (Liang *et al.*, 1998). A selective autophagy pathway, called xenophagy, acts as a defence mechanism in which microorganisms, such as bacteria, fungi, and viruses, are presented to the lysosomes for degradation (Virgin & Levine, 2009). In addition to directly degrading pathogens, upregulation of autophagy increases the innate immune response and prevents excessive inflammation.

Autophagy also plays a vital role in adaptive immunity (Levine *et al.*, 2011). The ATG5 gene is required for proper presentation of microbial antigens to major histocompatibility complex class II molecules (Lee, Mattei*, et al.*, 2010). Autophagy proteins are also involved in innate and adaptive immunity in an autophagy pathway-independent manner (English *et al.*, 2009, Levine *et al.*, 2011). Autophagy is a promising drug target to treat infectious disease, with some compounds already being tested in preclinical trials (Rubinsztein *et al.*, 2012).

1.3 The Unc51-like kinase 1 (ULK1) complex

Regulation of macroautophagy is well-studied in yeast (Tsukada & Ohsumi, 1993). It is initiated by the ATG1 complex, which consists of the Atg1, Atg13, Atg11 and Atg17 proteins (Kamada *et al.*, 2000). Upon starvation, the members of the ATG1 complex associate together and undergo a cascade of phosphorylation, which induces

macroautophagy (Kamada *et al.*, 2000). The mammalian homologue of the ATG1 complex, the Unc51-like kinase 1 (ULK1) complex, was identified initially by the Ohsumi lab in 1993 (Tsukada & Ohsumi, 1993), and later confirmed by the Tooze lab in 2006–2007 (Hara *et al.*, 2008, Chan *et al.*, 2007, Young *et al.*, 2006). The members of the ULK1 complex were first identified by co-immunoprecipitation experiments that identified focal adhesion kinase family-interacting protein of 200 kDa (FIP200) and Atg13 (Hara *et al.*, 2008, Hosokawa *et al.*, 2009). Like the ATG1 complex in yeast, the mammalian ULK1 complex consists of four proteins: ULK1, Atg13, FIP200 and Atg101 (Chan *et al.*, 2007).

Interestingly, the human ULK1 complex remains stably associated regardless of the activation or inactivation of autophagy. In contrast, the yeast ATG1 complex does not form stable complexes without the induction of autophagy (Kamada *et al.*, 2000, Kabeya *et al.*, 2005). Gel filtration chromatography experiments showed that ULK1 formed a stable 3-MDa complex in precipitates isolated from both starved and nutrient-rich mouse embryonic fibroblasts (Hosokawa *et al.*, 2009). These results were further confirmed using co-immunoprecipitation experiments (Hara *et al.*, 2008).

1.3.1 Cellular distribution

The cellular distribution of the ULK1 complex was first determined by expressing ULK1 proteins fused with green fluorescent protein (Hara *et al.*, 2008). Because macroautophagy can be induced by starvation (Mortimore & Poso, 1987), the cellular distribution of the ULK1 complex under starvation and nutrient-rich conditions was examined to verify the involvement of the ULK1 complex in macroautophagy. Under nutrient-rich conditions, the ULK1 protein was distributed

throughout the NIH/3T3 cells, whereas during starvation, the ULK1 protein was localised at punctate structures within the autophagosome, suggesting that the ULK1 protein was involved in macroautophagy (Hara *et al.*, 2008). Consistent with this result, both FIP200, which associates with ULK1, and Atg13 were spread throughout the cytoplasm under nutrient-rich conditions, but were localised near the isolation membrane and the punctate structure, respectively, during starvation (Hara *et al.*, 2008) (Hosokawa *et al.*, 2009). The localisation of these proteins to the autophagosome vesicle during starvation-induced macroautophagy suggests the involvement of ULK1-Atg13-FIP200 in macroautophagy.

1.3.2 Importance of the ULK1 complex in macroautophagy

To study which of the proteins in the ULK1 complex are necessary for autophagy, the genes encoding ULK1, Atg13 and FIP200 were mutated to produce loss-of-function variants (e.g. by mutating key residues involved in kinase activity) or complete knock-outs of individual genes. Autophagy was inhibited in kinase-dead ULK1 mutant cells (Hara *et al.*, 2008), and FIP200 knock-out mice showed embryonic lethality, caused by defective heart formation (Gan *et al.*, 2006). In addition, mouse embryonic fibroblast cells derived from these knock-out mice showed drastic suppression of macroautophagy (Hara *et al.*, 2008). Electron microscopy experiments demonstrated that GFP-tagged ULK1 was found near autophagosome punctates only in the wild type cells, and assembly was not observed in FIP200 knock-out cells (Hara *et al.*, 2008). These results suggest that FIP200 is essential in macroautophagy, and functions in complex with ULK1 (Hara *et al.*, 2008). Similar to ULK1 and FIP200, silencing of Atg13 using siRNA inhibits autophagosome formation (Hosokawa *et al.*, 2009).

1.3.3 Binding regions of the ULK1 complex

Binding assays suggest that FIP200 and Atg13 interact at the C-terminal region of ULK1 (Jung *et al.*, 2009), and ULK1 and FIP200 bind to the C-terminal region (384–517 aa) of Atg13 (Jung *et al.*, 2009). Although Atg13 interacts with ULK1 and FIP200 directly, *in vitro* studies suggest that ULK1 and FIP200 do not interact in the absence of Atg13 (Jung *et al.*, 2009). Thus, Atg13 mediates the interaction between FIP200 and ULK1 (Hosokawa *et al.*, 2009, Jung *et al.*, 2009).

1.4 Macroautophagy signalling pathways

Macroautophagy is well studied under starvation conditions, but is also induced by other environmental changes, such as energy depletion (Kim *et al.*, 2011), decreases in growth factors like interleukin-3 (Lum *et al.*, 2005), heat and oxidative stress (Scherz-Shouval *et al.*, 2007). There are various signalling pathways that regulate autophagy, which can be broadly classified as either mTOR-dependent or mTOR-independent (Metcalf *et al.*, 2012). The main focus of this thesis is the mTOR-dependent activation of the ULK1 complex.

1.4.1 mTOR-dependent signalling pathway

mTOR is a serine/threonine kinase that controls cell growth and protein synthesis by sensing energy and nutrients (Fleming *et al.*, 2011). The mammalian TOR exists in two complex forms: mTORC1 and mTORC2 (Laplante & Sabatini, 2012). The two complexes vary in their functions, as well as the subunits that they consist of. mTORC1 is defined by the association of regulatory-associated protein of mTOR complex 1 (RPTOR) subunit, whereas mTORC2 is defined by the association of RPTOR-independent companion of mTOR complex 2 (RICTOR) (Hara *et al.*,

2002, Dos *et al.*, 2004). Inactivation of mTORC1 leads to the activation of autophagy via the ULK1 complex (Hosokawa *et al.*, 2009). mTOR is activated by signals from insulin receptors that activate protein kinase B (PKB/Akt), which in turn phosphorylates the tuberous sclerosis complex (TSC1-TSC2 complex) (Huang & Manning, 2008), a heterodimer that acts as a GTPase-activating protein (Fleming *et al.*, 2011) (Figure 1.3). Co-immunoprecipitation experiments have shown that autophagy is negatively regulated by mTORC1, but not by mTORC2. In these studies, the RPTOR protein, which is part of mTORC1, interacted with ULK1, whereas components from mTORC2 did not interact with ULK1, suggesting the involvement of mTORC1, but not mTORC2, in autophagy (Hosokawa *et al.*, 2009).

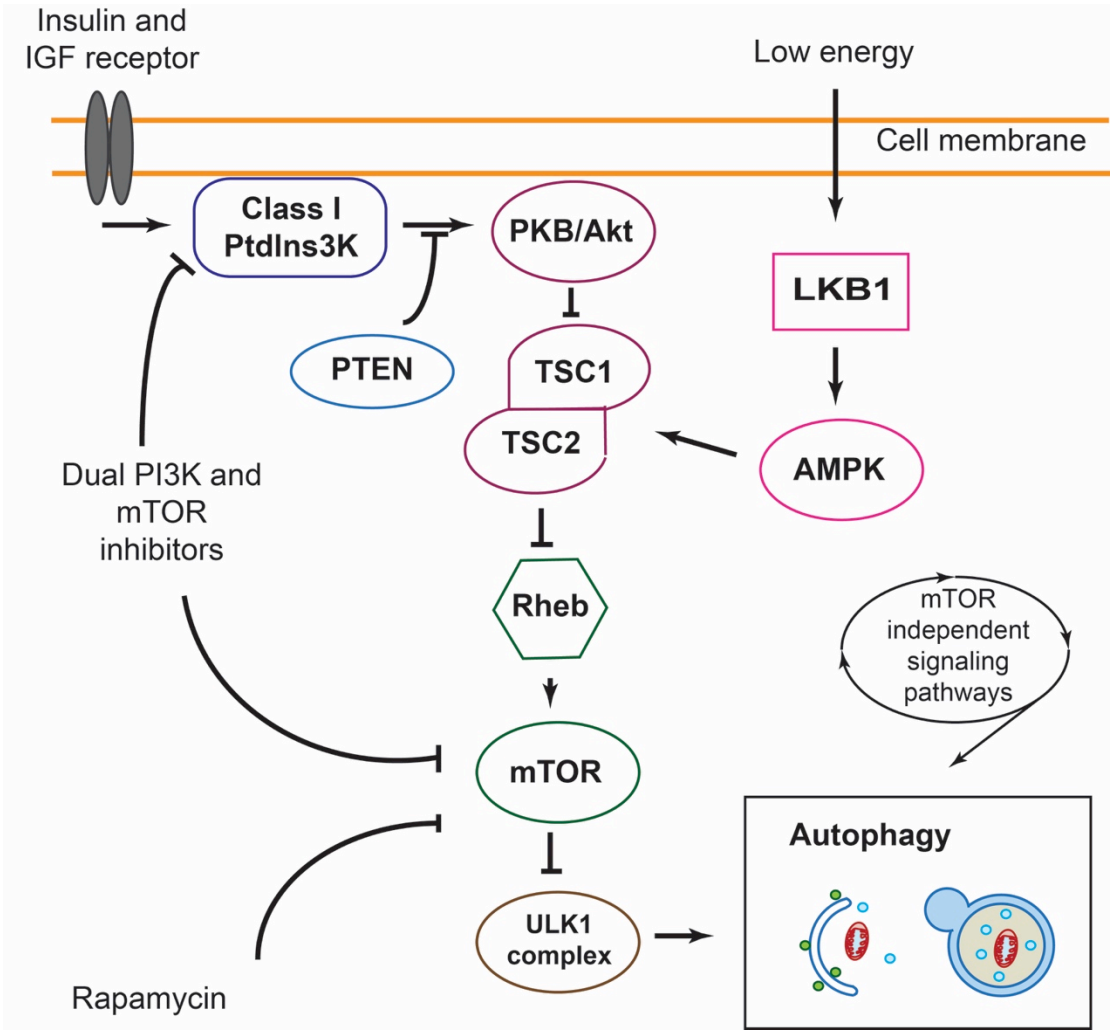


Figure 1.3: Schematic depicting the proteins involved in mTOR-dependent as well as mTOR-independent signalling pathways. mTOR is a well-studied starvation-induced regulator of the autophagy pathway that is activated by signals from insulin receptors. These signals activate protein kinase B (PKB/Akt), which in turn phosphorylates the TSC1-TSC2 complex. Inhibition of the TSC1-TSC2 complex activates mTOR, leading to the inhibition of autophagy. Adapted by permission from Macmillan Publishers Ltd: Nature Cell Biology, Fleming A., et al., December 16, 2010.

An *in vitro* kinase assay performed on a kinase-dead ULK1 mutant (to avoid autophosphorylation) from starved cells and mTORC1 from non-starved cells (Hosokawa *et al.*, 2009) suggested that ULK1 was phosphorylated when incubated with mTOR (Hosokawa *et al.*, 2009). Phosphorylation of Atg13 was also detected, showing that mTOR phosphorylates both ULK1 and Atg13. However, Atg13 was still phosphorylated when ULK1 was silenced, suggesting that Atg13 could be a direct substrate for both mTOR and ULK1 (Hosokawa *et al.*, 2009). Similar studies

performed on FIP200 and ULK1 revealed that ULK1 phosphorylates FIP200 in an Atg13-dependent manner (Hosokawa *et al.*, 2009, Jung Ch Fau - Jun *et al.*, 2009). Under starvation, mTOR dissociates from the ULK1 complex (Hosokawa *et al.*, 2009), leading to the conclusion that inhibition of mTOR prevents the phosphorylation of ULK1 and Atg13, which subsequently causes ULK1 to phosphorylate FIP200 (Jung Ch Fau - Jun *et al.*, 2009).

In the mTOR-dependent pathway, energy depletion or nutrient stress activates AMP-activated protein kinase, which in turn inactivates mTORC1 and negatively regulates autophagy. Insulin and other growth factor signals also inactivate mTORC1 (Metcalf *et al.*, 2012). In contrast, the mTOR-independent signalling pathways are regulated by calcium, cyclic AMP, calpains and inositol triphosphate receptor (Metcalf *et al.*, 2012).

1.5 Selective autophagy

Autophagy was originally considered a non-selective pathway for bulk degradation of protein. However, recent evidence shows that it can selectively bind to a cargo and degrade it (Johansen & Lamark, 2011). Selective autophagy was first identified in yeast, which use the cytoplasm-to-vacuole pathway (Cvt) to selectively degrade aminopeptidase (Ape1) and α-mannosidase (Ams1) (Harding *et al.*, 1995, Hutchins & Klionsky, 2001). After recognition of the cargo by a receptor, adaptor proteins link the cargo to the main autophagic machinery for degradation (Johansen & Lamark, 2011). Different cargo-dependent terms for autophagy are described in Table 1.1, and are shown in Figure 1.4. The ubiquitin-binding protein p62 is a well-studied selective autophagy receptor that binds to ubiquitinated protein aggregates, bacteria, peroxisomes and post-mitotic mid-body ring structures and links them to the

autophagy machinery (Ichimura, Kumanomidou, *et al.*, 2008). Other examples of selective autophagy receptors include NBR1 (Kirkin *et al.*, 2009), NDP52 (Thurston *et al.*, 2009) and Nix (Novak *et al.*, 2010). These receptors have a cargo-recognition domain containing a specific motif, called the LC3 interacting region (LIR), which is recognised by the mammalian Atg8 protein family (LC3A, LC3B, LC3C, GABARAP, GABARAPL1, GABARAPL2) (Kirkin *et al.*, 2009, Novak *et al.*, 2010, Thurston *et al.*, 2009). The LIR interaction is discussed in Section 1.6.1.

Table 1.1: **Types of selective autophagy and their definitions**.

Type of selective Autophagy	Definition
Aggrephagy	Selective autophagic sequestration of protein aggregates.
Crinophagy	Direct fusion of secretory vesicles with lysosomes.
Cytoplasm to vacuole targeting (Cvt)	A biosynthetic pathway in yeast that transports resident hydrolases to the vacuole (the yeast lysosome) through a selective autophagy-related process.
Mitophagy	Selective autophagic sequestration and degradation of mitochondria.
Pexophagy	Selective autophagy involving the sequestration and degradation of peroxisomes; can occur by a micro- or macropexophagic process.
Piecemeal microautophagy of the nucleus	Intrusion of portions of the nucleus into the vacuole, followed by scission and degradation.
Reticulophagy	Selective autophagic sequestration and degradation of endoplasmic reticulum.
Vacuole import and degradation	Selective uptake of cytosolic fructose-1,6-bisphosphatase, and possibly other proteins, within 30-nm single membrane vesicles, followed by fusion with the vacuole and degradation.
Xenophagy	Selective degradation of microbes (e.g., bacteria, fungi, parasites and/or viruses) through an autophagy-related mechanism.

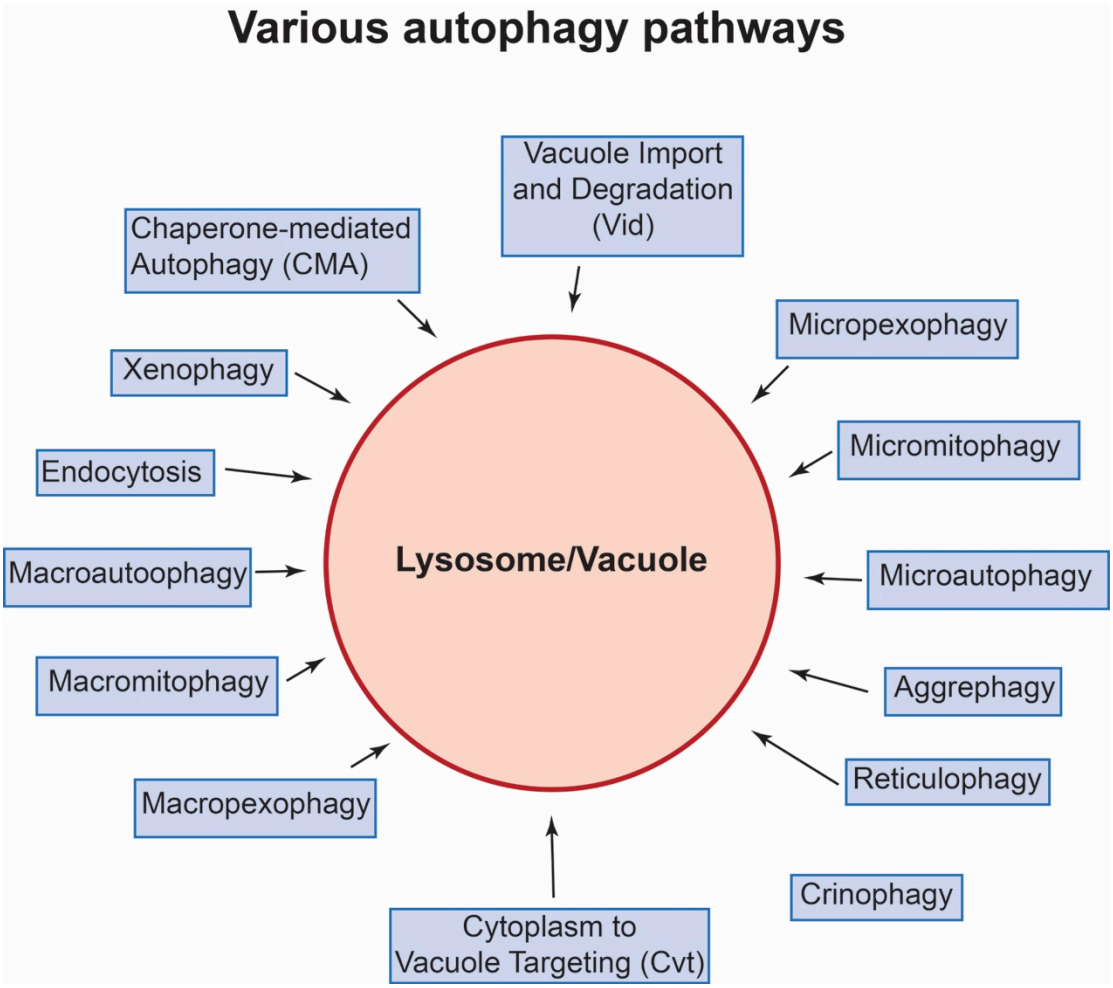


Figure 1.4: Figure depicting the various selective autophagy pathways. All of the autophagy pathways shown result in fusion with the lysosome, which degrades the cargo and unwanted cytosolic material carried by the vesicles.

1.6 Atg8 family proteins

Atg8 is a ubiquitin-like protein that was first identified in yeast and is required for autophagosome formation (Nakatogawa *et al.*, 2007). Conjugation of Atg8 with phosphatidylethanolamine (PE) is an important event in autophagosome formation (Kabeya *et al.*, 2000). Atg4, a cysteine protease, processes Atg8 by removing the C-terminal residue, leaving a glycine at the newly formed C-terminus (Kabeya *et al.*, 2000, Kabeya *et al.*, 2004, Tanida *et al.*, 2004). Atg7 and Atg3, which are similar to

E1-like activating and E2-like conjugating enzymes, respectively, conjugate the C-terminus of Atg8 to PE (Kabeya *et al.*, 2000, Kabeya *et al.*, 2004). PE-conjugated Atg8 localises to the phagophore assembly site and helps in autophagosome formation (Nakatogawa *et al.*, 2007). It also controls the size of the autophagosome and the level of autophagy in cells (Mizushima *et al.*, 2011, Xie *et al.*, 2008).

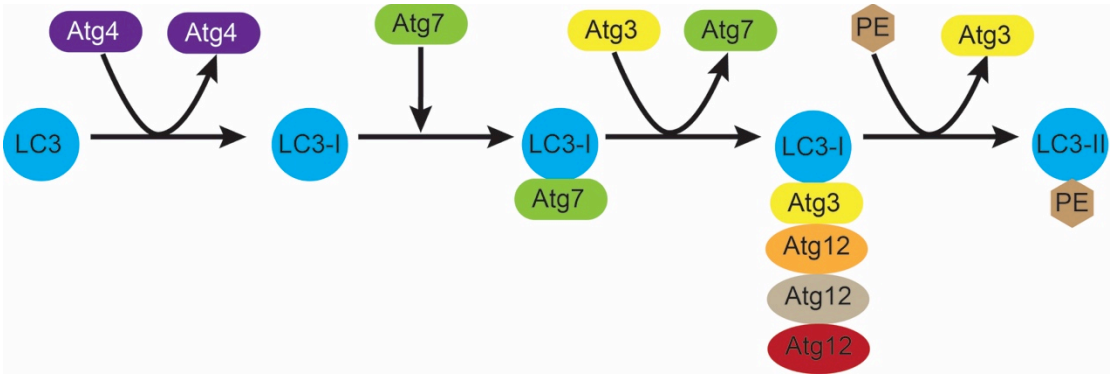


Figure 1.5: Various steps involved in the conjugation of phosphatidylethanolamine (PE) to LC3-I. First, the C-terminal of LC3 is cleaved by Atg4 to form LC3-I, with an exposed C-terminal glycine. Next, Atg7 and Atg3, which are similar to E1 and E2 ligase, conjugate the PE to LC3-II.

Microtubule-associated protein I light chain 3 (LC3) was identified as a mammalian homologue of Atg8, with 28% identity shared between the two proteins (Mann & Hammarback, 1994). At least eight Atg8 homologues have been identified in mammals (He *et al.*, 2003, Igloi, 2001), and can be divided into two subgroups based on sequence homology: the LC3 isoform subgroup, consisting of LC3A, LC3B and LC3C, and the GABARAP isoform subgroup, consisting of GABARAP, GABARAPL1, GABARAPL2 and GABARAPL3 (He *et al.*, 2003, Igloi, 2001). The LC3 isoforms exist as LC3-I (found in cytosol) and LC3-II (bound to the autophagosome membrane) (Sou *et al.*, 2006). Similar to yeast Atg8, cleavage of LC3 isoforms at the C-terminus forms LC3-I, which also has a glycine at the newly formed C-terminus (Figure 1.5). Mammalian Atg7 (E1-like enzyme) forms a thioester bond

between its activated cysteine residue and the C-terminal glycine of LC3-I, which is then transferred to mammalian Atg3 (E2-like ligase) (Figure 1.5). Mammalian Atg3 recognises PE and conjugates it to LC3-I, forming LC3-II (LC3-PE) (Figure 1.5) (Kabeya *et al.*, 2000, Sou *et al.*, 2006). Because there are several orthologues of Atg8 in mammals, the question arises of whether they have distinct roles in autophagy (Figure 1.6).

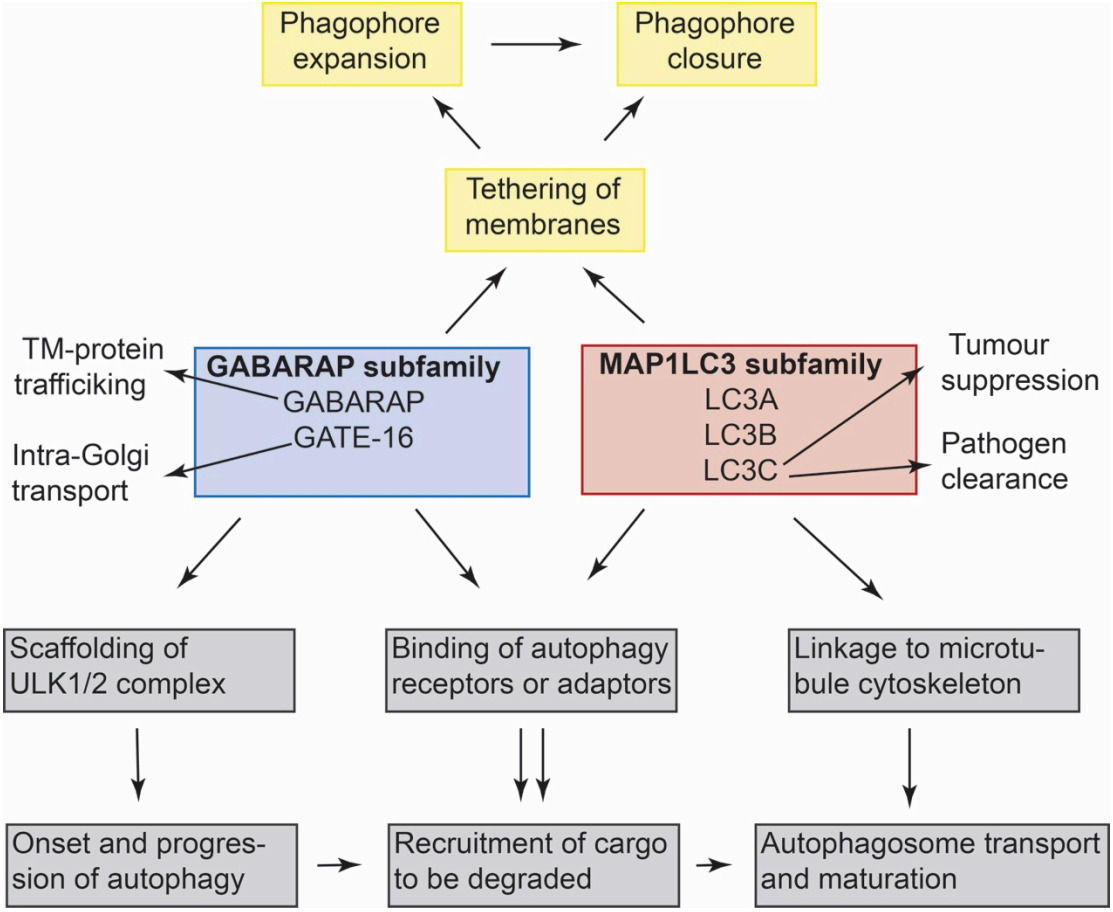


Figure 1.6: Various functions of the LC3 and GABARAP subfamilies of mammalian Atg8 orthologues. The LC3 family is involved at the beginning of autophagosome formation, whereas the GABARAP family act at the autophagosome closure step. © 2013 Weiergräber OH, Mohrlüder J, Willbold D. under CC BY 3.0 license. Available from: http://dx.doi.org/10.5772/55647.

Studies by Weidberg *et al*. (2010) suggest that both the LC3 and GABARAP subfamilies promote autophagosome biogenesis, but that they act at different stages of the process. The LC3 subfamily is more involved at the beginning of autophagosome

formation, whereas the GABARAP subfamily is involved during the closure step of autophagosome formation (Weidberg *et al.*, 2010).

1.6.1 LIR interaction

Atg8 family proteins are the adaptors that link the cargo-recognizing receptors to the core autophagy machinery (Ichimura, Kumanomidou*, et al.*, 2008, Pankiv *et al.*, 2007). In addition to this role, LC3 and GABARAP family proteins also act as scaffolds for the ULK1 complex (Figure 1.6) (Alemu *et al.*, 2012). This requires the interaction of ATG proteins with LC3s and GABARAPs through the specific LIR motif (Alemu *et al.*, 2012, Pankiv *et al.*, 2007). A common core sequence of [**W/F/Y**]x_1x_2[**L/I/V**] (where x = any amino acid) in the LIR motif was identified from crystal structures of cargo receptors in complex with Atg8 family protein, and by deletion and point mutation analysis of p62 and Atg19 (an adaptor in the Cvt pathway) (Ichimura, Kumanomidou*, et al.*, 2008, Kirkin *et al.*, 2009, Noda *et al.*, 2008, Novak *et al.*, 2010, Pankiv *et al.*, 2007, Rogov *et al.*, 2013, Suzuki *et al.*, 2014, Thurston *et al.*, 2009, Lystad *et al.*, 2014). Various LIR-containing proteins and their LIR motifs are shown in Table 1.2.

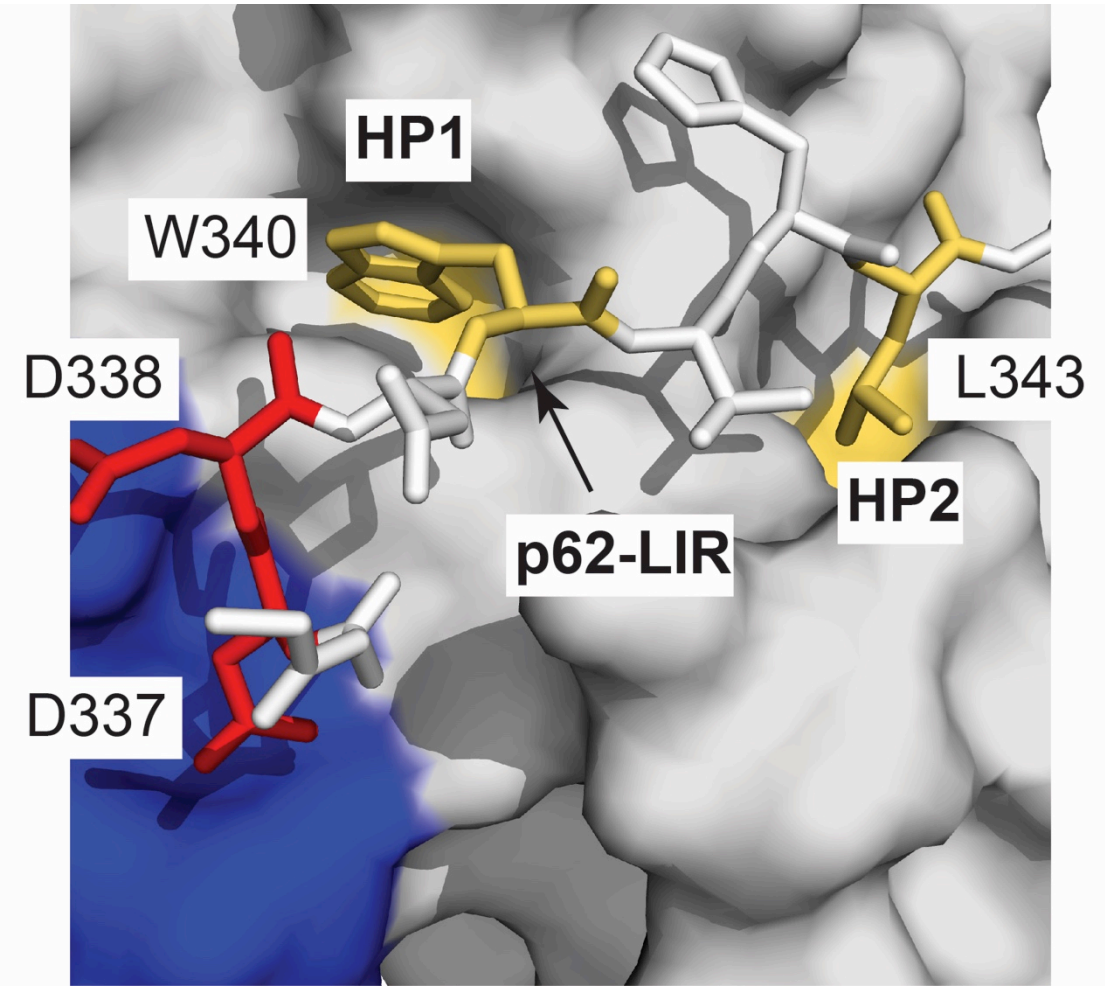


Figure 1.7: LC3B crystal structure in surface representation bound to the p62 LIR peptide. The hydrophobic pockets HP1 and HP2 are shown as yellow surfaces, and the LIR amino acids that interact with these binding pockets are shown as yellow sticks. The acidic residues (red) in the LIR peptide interact with a basic (blue) region of LC3B.

The LIR motif consists of a core of hydrophobic residues in which the N-terminal residue is usually aromatic (Tyr, Phe or Trp) (Alemu *et al.*, 2012). Mutational studies on these core aromatic residues showed that Trp had higher binding affinity than Phe and Tyr (Rozenknop *et al.*, 2011). The N-terminal aromatic residue sits into hydrophobic pocket 1 (HP1) of the LIR docking site in Atg8 proteins (Ichimura, Kumanomidou, *et al.*, 2008, Pankiv *et al.*, 2007). The C-terminal hydrophobic residue (leucine, isoleucine or valine) binds to hydrophobic pocket 2 (HP2) (Figure 1.7) (Ichimura, Kumanomidou, *et al.*, 2008, Pankiv *et al.*, 2007). LIRs containing phenylalanine display a preference for specific amino acid residues at the

"x_1" position (Noda *et al.*, 2008, Ichimura, Kominami*, et al.*, 2008), with mutational studies showing that only valine, cysteine, isoleucine and glutamic acid are acceptable at this position (Alemu *et al.*, 2012).

Some LIRs are specific to particular Atg8 family members. For example, the NDP52 LIR is specific to LC3C (von Muhlinen *et al.*, 2012a). This LIR does not have an N-terminal aromatic residue, which means it is unable to bind to other Atg8 proteins (von Muhlinen *et al.*, 2012a). However, a structural rearrangement of the LIR creates an interaction with LC3C sequence-specific residues (von Muhlinen *et al.*, 2012b). Acidic residues surround the LIR motif and play an important role in forming polar interactions at the LIR docking site, which was verified by alanine screening (Noda *et al.*, 2008, Pankiv *et al.*, 2007). Thus, LIR interaction is an important part of autophagy pathways.

Table 1.2: Different LIR motif-containing binding partners of mammalian Atg8 family proteins.

Protein	LIR motif	Pos.
ULK1	SCDTDD**F**VM**V**PA	357
ULK2	SCDTDD**F**VL**V**PH	353
DmATG1B	HEDSDD**F**VL**V**PK	391
ScAtg1	RSFERE**Y**VV**V**EK	391
ATG13	GNTHDD**F**VM**I**DF	429
FIP200	DAHTFD**F**ET**I**PH	702
p62	SGGDDD**W**TH**L**SS	338
NBR1	SASSED**Y**II**I**LP	732
AtNBR1	LCGVSE**W**DP**I**LE	661
ATG4B	DAATLT**Y**DT**L**RF	8
TP53INP1	EKEDDE**W**IL**V**DF	31
TP53INP2/DOR	EDEVDG**W**LI**I**DL	36
FYCO1	PPDDAV**F**DI**I**TD	1280
Dv12	EVRDRM**W**LK**I**TI	444
c-Cb1	ASSSFG**W**LS**L**DG	802
NIX	AGLNSS**W**VE**L**PM	36
FUNDC1	ESDDDS**Y**EV**L**DL	18
Calreticulin	GSLEDD**W**DF**L**PP	183
Clathrin HC	VGYTPD**W**IF**L**LR	513
OATL1/TBC1D25	SPLLED**W**DI**I**SP	136
TBC1D5 LIR2	SSKDSG**F**TI**V**SP	788
Optineurin	GSSEDS**F**VE**I**RM	178
Stbd1	RVDHEE**W**EM**V**PR	203
ScAtg3	LDGVGD**W**ED**L**QD	270
ScAtg19	NEKALT**W**EE**L**	412
ScAtg32	DSISGS**W**QA**I**QP	86

1.7 Aims of this project

The aim of this thesis was to study the interactions between the proteins involved in macroautophagy. These interactions can be broadly classified into LIR-dependent and LIR-independent interactions (Behrends *et al.*, 2010). The interaction between FIP200 and Atg13, which are components of the ULK1 initiation complex, was probed as an example of LIR-independent interactions. As mentioned in Section 1.3.3, the region of Atg13 (C-terminal 384–517) that interacts with FIP200 is known, whereas the Atg13-binding region of FIP200 is unknown. To map the Atg13-binding region on FIP200, the full-length FIP200 gene was divided into six smaller constructs, and pull-down assays were performed to identify the FIP200 construct that binds to the C-terminal region of Atg13.

To study LIR-dependent interactions, two LIR motifs were chosen: pleckstrin homology domain-containing protein family member 1 (PLEKHM1) LIR and ULK1 LIR. As mentioned in Section 1.6, there are six mammalian Atg8 homologues, as compared to one Atg8 protein in yeast. Unpublished data from our collaborators demonstrated that all six members of the mammalian Atg8 family have different binding affinities towards LIR peptides, with GABARAP family proteins demonstrating significantly higher affinity than LC3 family proteins. In this thesis, determination of the crystal structures of all six human Atg8 proteins in complex with the PLEKHM1 LIR was attempted to gain insight into the reasons for the different binding affinities.

To study the differences in the mechanism by which LC3 and GABARAP family members interact with LIRs, the interaction of the ULK1-LIR peptide with LC3A and GABARAP was probed using surface plasmon resonance (SPR), X-ray crystallography and molecular dynamics simulations.

Crystal structures of LC3s with Atg13 LIRs demonstrated Lys49 in LC3s undergoes structural rearrangement upon LIR binding and has a significant biological function. Mutating Lys49 to Ala reduces autophagosome formation (Suzuki *et al.*, 2014). In this thesis, the significance of Arg70 and Phe52 in LC3A, and Arg67 and Tyr49 in GABARAP were studied using SPR. Overall, this thesis aims to shed light on the interactions that contribute to the proper functioning of macroautophagy.

Chapter 2. EXPERIMENTAL

2.1 Materials and equipment

2.1.1 Reagents

Unless otherwise stated, chemicals were purchased from Sigma-Aldrich Chemicals Ltd. (St Louis, MO, USA). Media for growing bacterial cultures were purchased from Life Technologies Ltd. (Christchurch, New Zealand). General reagents used are summarised in Table 2.1. Solutions were prepared using purified water from a MilliQ (MQ) system (Millipore, Billerica, MA, USA). Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels and protein ladders were purchased from Life Technologies Ltd. Column chromatography media were purchased from GE Healthcare Life Sciences (Little Chalfont, UK). Centrifugation was performed using Eppendorf 5810R (fixed-angle rotor F-34-6-38; Hamburg, Germany) and Sorvall RC 6 plus (rotor F10S6×500Y; ThermoFisher Scientific, Waltham, MA, USA) centrifuges.

Table 2.1: **Summary of general reagents used in this work**

Item	Supplier
Acetic acid, glacial	Ajax Finechem
Agar (bacteriological)	Life Technologies
Ampicillin	Life Technologies
Ammonium sulphate	Astral
Benchmark protein ladder	Life Technologies
Bromophenol blue	Progen
Calcium chloride	BDH
Chloramphenicol	BDH
Coomassie brilliant blue	BDH
Dimethylformamide	Astral
Dithiothreitol	BDH
DNase I	Astral
Glucose	Astral
Glycerol	AppliChem
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid	Astral
Hydrochloric acid	Ajax Finechem
Imidazole	AppliChem
Isopropyl-β-D-thiogalactopyranoside (IPTG)	Astral
Lactose	Merck
Lysozyme	Astral
Magnesium sulphate	AppliChem
Potassium chloride	BDH
Potassium phosphate	Astral
Protease inhibitor cocktail	Sigma-Aldrich
RNase A	Astral
Sodium acetate	Ajax Finechem
Sodium chloride	Ajax Finechem
Sodium dodecyl sulphate	AppliChem
Sodium hydroxide pellets	BDH
Sodium phosphate	Astral
Tetramethylethylenediamine	Bio-Rad
Tris(hydroxymethyl)aminomethane	Astral
Bacto-tryptone	Oxoid
Urea	Merck
Yeast extract	Life Technologies

2.1.2 Growth media and buffers

2.1.2.1 Luria-Bertani (LB) medium

Bacto-tryptone (10 g), yeast extract (5 g) and NaCl (10 g) were added to 1 L of MQ water and sterilised by autoclaving. LB plates were prepared by including 15 g/L agar prior to autoclaving.

2.1.2.2 SOC medium

Bacto-tryptone (20 g), yeast extract (5 g), NaCl (0.58 g), KCl (0.18 g), $MgCl_2$ (0.95 g) and $MgSO_4$ (1.20 g) were added to 1 L of MQ water and sterilised by autoclaving. Filter-sterilised glucose (20% w/v) was then added.

2.1.2.3 LB agar plate preparation

LB agar plates were prepared by adding 1.5% w/v agar to LB media (Section 2.1.2.1) prior to autoclaving. After LB agar had cooled, but was still in the molten state, antibiotics were added and the medium was poured into plates, which were then allowed to cool.

2.1.2.4 Lysis buffer

A buffer with final concentrations of 50 mM Tris-HCl (pH 9.0), 250 mM NaCl and 10 mM imidazole was prepared in MQ water.

2.1.2.5 Phosphate buffered saline (PBS)

Five PBS tablets (Sigma-Aldrich) were dissolved in 1 L of MQ water to give a buffer containing 0.01 M phosphate buffer, 2.7 mM KCl, and 137 mM NaCl. The pH was altered to either 7.2 or 7.4, as required. The buffer was filtered using a 0.4-μM

filter and stored at 4°C.

2.1.2.6 Elution buffer

A buffer with final concentrations of 50 mM Tris-HCl (pH 8.0), 250 mM NaCl and 500 mM imidazole was prepared in MQ water.

2.1.2.7 Size exclusion buffer

A buffer with final concentrations of 20 mM Tris-HCl (pH 8.0) and 100 mM NaCl was prepared using MQ water.

2.1.3 Expression hosts and plasmid

The bacterial strains and mammalian cells used in this work are listed in Table 2.2.

Table 2.2: Summary of bacterial strains and mammalian cells used in this work

***Escherichia coli* strains**	**Use**	**Source**
BL21 (DE3)	expression host	Life Technologies
Rosetta (DE3)	expression host	Life Technologies
BL21 (DE3) pLysS	expression host	Life Technologies
DH5α	plasmid isolation	Life Technologies
XL1-Blue	plasmid isolation	Life Technologies
Mammalian cells	**Use**	**Source**
FreeStyle 293-F Cells	expression host	Life Technologies

Antibiotics used in the media were filter-sterilised and stored as 1000× stock solutions at −20°C (Table 2.3). The final working concentrations are those used to select for appropriate bacterial strains.

Table 2.3: **Antibiotic concentrations.**

Antibiotic	Stock (mg/mL)	Working (µg/mL)
Ampicillin	100	100
Chloramphenicol	30	30
Kanamycin	50	50

Sequences encoding the LC3 [LC3A (accession: NP_115903), LC3B (accession: NP_073729), and LC3C (accession: NP_001004343)] and GABARAP [GABARAP (accession: CAG33324), GABARAPL1 (accession: CAG38511), and GABARAPL2 (accession: CAG47013)] family proteins were cloned into vectors pET30delSE and pGEX-4T-1 (GE Healthcare Life Sciences). All the amino acids mentioned in this thesis were numbered according to the numbering scheme followed in National Centre for Biotechnology Information (NCBI) database. The mutants LC3A-R70L, LC3A-F52Y, GABARAP-R67L and GABARAP-Y49F were cloned into pET30delSE. Chimeric constructs of the PLEKHM1-LIR fused to the N-terminus of LC3A, GABARAP and GABARAPL1, and the ULK1-LIR fused to the N-terminus of GABARAP, were cloned into pET30delSE. The Atg13 and ULK1 constructs were cloned into pGEX-4T-1. FIP200 constructs were cloned into pCOLD TF (Clontech, Mountain View, CA, USA). The full-length genes of GABARAP, GABARAPL1 and GABARAPL2 were kindly provided by Prof. Ivan Dikic, Goethe-University, Frankfurt, Germany. Dr Hironori Suzuki, a previous post-doc in our lab, kindly provided all the other full-length genes used in this work. A summary of all the proteins used in this work is shown in Table 2.4.

Table 2.4: **Summary of all the proteins used in this work**

Protein	Number of amino acids
ATG13 (Full length)	517
ATG13 380–517	137
FIP200 2–265 (P1)	263
FIP200 266–531 (P2)	265
FIP200 532–796 (P3)	264
FIP200 797–1061 (P4)	264
FIP200 1062–1326 (P5)	264
FIP200 1327–1591 (P6)	264
LC3A 2–121	119
LC3B 2–124	122
LC3C 8–125	117
GABARAP 2–117	115
GABARAPL1 2–117	115
GABARAPL2 2–117	115
LC3A-R70L	119
LC3A-F52Y	119
GABARAP-R67L	115
GABARAP-Y49F	115
ULK1LIR-GABARAP	121
PLEKHM1LIR-LC3A	125
PLEKHM1LIR-GABARAP	123
PLEKHM1LIR-GABARAPL1	123

2.1.4 Chromatography equipment, columns and media

Chromatography was performed on an ÄKTApurifier system (GE Healthcare). Hi-Trap chelating columns (1 mL, GE Healthcare) and HisTrap columns (1 mL, GE Healthcare) were used for immobilised metal affinity chromatography (IMAC), precharged with different metal ions (Section 2.5.2). Superdex 200 GL 10/300 (GE Healthcare) columns (24 mL) and HiLoad 16/600 Superdex 200 pg (GE Healthcare)

columns (120 mL) were employed for analytical and preparative size exclusion chromatography (SEC), respectively.

2.2 Molecular biology

2.2.1 Isolation of plasmids

Using aseptic technique, a single colony of transformed DH5α cells was picked from a transformation plate using an autoclaved pipette tip, and then transferred to 10 mL of sterile LB medium containing the appropriate antibiotic for plasmid selection. Cells were cultured overnight at 37°C with shaking at 180 rpm in a Minitron incubator (Infors HT, Basel, Switzerland). Cells from a 5-mL aliquot of overnight culture were then harvested by centrifugation at 4°C for 5 min at 10,000 rpm. The plasmid was extracted from the cells and purified using a Purelink® Quick Plasmid Miniprep Kit (Invitrogen™ by Life Technologies™, Carlsbad, CA, USA) according to the manufacturer's instructions. Plasmid DNA was eluted in 50 μL of Tris-EDTA (TE) buffer (supplied with kit), and the concentration of plasmid was checked using a NanoDrop ND-1000 UV-VIS spectrophotometer (ThermoFisher Scientific) at 260 nm. All purified plasmids were stored at −20°C.

2.2.2 Cloning

2.2.2.1 Polymerase chain reaction (PCR)

DNA amplification was carried out using KOD Hot Start DNA Polymerase® (Novagen® by Merck Millipore, Billerica, MA, USA). A 10× buffer for KOD Hot Start DNA Polymerase, 2 mM dNTPs and 25 mM $MgSO_4$ were provided in the KOD

Hot Start DNA Polymerase® kit. PCR reactions were set up in PCR tubes according to Table 2.5.

Table 2.5: **PCR reaction set up**

PCR components	**Volume**
DNA template	1 µL
10× PCR buffer	2.5 µL
$MgSO_4$	1.5 µL
dNTPs	2.5 µL
Forward primer	0.75 µL
Reverse primer	0.75 µL
KOD Hot Start DNA Polymerase®	0.5 µL
Autoclaved MQ water	15.5 µL
Total volume	**25 µL**

PCR assays were performed using a MultiGene™ PCR Thermal Cycler (LabNet International Inc., Edison, NJ, USA), using the cycling parameters in Table 2.6.

Table 2.6: **PCR cycling parameters.**

STEP	**TIME**	**TEMPERATURE**	**CYCLES**
Initial denaturation	2 min	94 °C	1
Denaturation	1 min	94 °C	
Annealing	1 min	53 °C	25
Extension	1.5 min	72 °C	
Final extension	7 min	72 °C	1

A final extension time of 7 min at 72°C was included to ensure all PCR products were full-length and 3′ adenylated (required for TA cloning). Following the final cycle of PCR amplification, the reaction was maintained at 4°C. PCR amplification products were analysed by DNA electrophoresis by loading 5 µL of

each PCR reaction with DNA loading dye into each well of a 1% w/v agarose gel (Section 2.2.3).

2.2.2.2 Sequence insertion for chimeric constructs

Chimeric constructs were prepared by fusing a sequence encoding a peptide to the N-terminal of the gene of interest in the expression vector. The oligonucleotide encoding the peptide region (ULK1 LIR or PLEKHM1 LIR) with restriction sites at each end was annealed with its complementary sequence using a MultiGene™ PCR Thermal Cycler. The oligonucleotides were dissolved in autoclaved MQ water to obtain 100 µM stock solutions of each. The reaction mix was prepared in a PCR tube by adding 1 µL of each oligonucleotide, 1 µL of 5 M NaCl and 17 µL MQ water. Thermal cycling parameters are provided in Table 2.7.

Table 2.7: Chimera construct cycling parameters

TEMPERATURE	TIME (min)
98°C	2
80°C	3
70°C	4
65°C	5
60°C	6
55°C	7
50°C	8
45°C	9
4°C	Hold

Following restriction enzyme digestion (Section 2.2.4) of the annealed PCR product, it was ligated (Section 2.2.4) into the target vector used for preparing the chimeric construct. The presence of insert was verified by colony PCR (Section 2.3.2.4) and by DNA sequencing (Section 2.2.6).

2.2.2.3 Site-directed mutagenesis

An inverse PCR technique was used to introduce site-directed point mutations. A forward primer was designed such that the 5′ end of the primer encoded the desired mutant amino acid, followed by 20–25 bp corresponding to the template sequence. The reverse primer contained 20–25 bp of template sequence, with the 3′-most three base pairs corresponding to the desired mutant amino acid sequence. After designing the primers, the PCR assay was carried out according to Table 2.8.

Table 2.8: Mutagenesis PCR cycling parameters.

TEMPRATURE	TIME	CYCLE
98°C	2 min	
98°C	15 s	
72°C	1 kb/min	18
72°C	7 min	
4°C	Storage	

Following PCR, 2.5 µL of PCR product were used for ligation. Because the PCR produced blunt ends, 0.5 µL of T4 DNA kinase was added to the ligation mixture and incubated at 37°C for 1 h. The plasmid was isolated and the sequence containing the mutation was verified by sequencing.

2.2.2.4 Colony PCR

Colony PCR of *Escherichia coli* DH5α colonies produced as a result of ligation and transformation was used to check for the presence and correct size of the insert within the plasmid of interest. Using aseptic technique, each colony was picked from an LB agar plate using an autoclaved toothpick and resuspended in 20 µL of sterile MQ water. Up to 10 colonies per plate were randomly picked and screened for the correct insert. PCR amplification was performed using GoTaq® Green Master Mix (Promega Corporation, Madison, WI, USA) in reactions set up according to

Table 2.9.

Table 2.9: **Colony PCR reaction set up.**

PCR components	**Volume**
Resuspended colony	2 µL
Forward primer	0.5 µL
Reverse primer	0.5 µL
GoTaq® Green Master Mix	7.5 µL
Autoclaved MQ water	4.5 µL
Total Volume	**15 µL**

PCR assays were carried out using a MultiGene™ PCR Thermal Cycler, according to the cycling parameters in Table 2.10.

Table 2.10: Colony PCR reaction cycling parameters.

STEP	**TIME**	**TEMPERATURE**	**CYCLES**
Initial denaturation	3 min	95°C	1
Denaturation	45 s	95°C	
Annealing	15 s	53°C	30
Extension	1.5 min	72°C	
Final extension	5 min	72°C	1

Following the final cycle of PCR amplification, the reaction was maintained at 4°C. PCR amplification products were analysed by DNA electrophoresis by loading the entire 15 µL colony PCR reaction mixture into wells of a 1% w/v agarose gel (Section 2.2.3).

The remaining 18 µL of each resuspended colony were stored at 4°C. Colonies containing a plasmid with the correct sized DNA insert were added to 10 mL of sterile LB medium containing the appropriate antibiotic. Cells were cultured overnight at 37°C, with shaking at 180 rpm, in a Minitron incubator prior to plasmid purification (Section 2.2.1).

2.2.3 Agarose gel analysis

Agarose gels were prepared by adding 0.5 g of agarose to 50 mL of 1× Tris acetate EDTA (TAE) buffer (see recipe in Appendix), and microwaving the solution for approximately 2 min until the agarose was completely dissolved. SYBR® Safe DNA Gel Stain (3 µL; Invitrogen™) was added to the dissolved agarose solution and swirled to mix, to allow DNA visualisation following electrophoresis. The 1% agarose/TAE solution was then poured into a gel casting box, and a comb was inserted to form wells. The gel was allowed to set at room temperature.

Once set, the agarose gel was removed from the gel casting box and placed in a Mini Gel Unit (Hoefer Inc., San Francisco, CA, USA), and 1× TAE buffer was added until the gel was covered. DNA loading dye was added to a final concentration of 1× (samples containing GoTaq® Green Master Mix (Promega Corporation) required no loading dye) and, once mixed, were loaded into wells. HyperLadder™ 1 kb (3 µL; Bioline USA Inc., Taunton, MA, USA) was loaded into one of the wells as a molecular weight marker. DNA gels were run for 30 min at 160 V using a Bio-Rad PowerPac Basic (Bio-Rad Laboratories, Hercules, CA, USA).

Following electrophoresis, DNA was visualised using a Chemi Genius 2 Bio Imaging System and GeneSnap software (SynGene, Cambridge, UK) using transilluminator darkroom lighting and the EtBr/UV filter.

2.2.4 Restriction and ligation

All of the restriction enzymes used in this thesis work were High-Fidelity® restriction enzymes from New England Biolabs (Ipswich, MA, USA). Reactions were performed in the supplied CutSmart™ buffer. Restriction reactions were set up as in Table 2.11.

Table 2.11: Restriction reaction composition.

RESTRICTION COMPONENTS	VOLUME
Plasmid	5 µL
Restriction enzyme I	1 µL
Restriction enzyme II	1 µL
CutSmart™ Buffer	2 µL
Autoclaved MQ water	11 µL
Total	**20 µL**

Reaction mixtures were incubated for 1 h at 37°C, and the restriction enzyme was deactivated either by heating at 65°C or using a PCR product clean up kit (Qiagen, Hilden, Germany).

Ligations were performed using DNA Ligation Kit Mighty Mix (Takara Clontech Laboratories, Mountain View, CA, USA). Reactions were set up as shown in Table 2.12.

Table 2.12: Ligation reaction composition.

LIGATION COMPONENTS	VOLUME
Vector	1 µL
Insert	1.5 µL
Ligase	2.5 µL
Total	**5 µL**

The reaction mixture was incubated for 15–20 min at room temperature, and then the entire mixture was transformed into chemically competent cells (Section 2.2.5).

2.2.5 Competent cell preparation

A 10-mL starter culture (SOC medium) was grown overnight at 37°C with shaking at 180 rpm. This starter culture was used to inoculate 300 mL of LB medium

without antibiotics. The culture was incubated at 37°C, with shaking at 180 rpm, until the optical density at 600 nm (OD_{600}) reached 0.6. The cells were centrifuged at 6000 rpm for 10 min at 4°C, and the supernatant was discarded. The cells were resuspended in 30 mL (1/10 of the culture volume) of ice-cold 0.1 M $CaCl_2$ and incubated on ice for 4 h. The $CaCl_2$ was removed by centrifuging the cell suspension at 6000 rpm for 10 min at 4°C and discarding the supernatant. The cells were again resuspended in 7.5 mL (1/40 of culture volume) of 0.085 M $CaCl_2$/15% glycerol solution. Aliquots of 50 μL were prepared and were either used immediately or snap frozen in liquid nitrogen and stored at −80°C.

2.2.6 DNA sequencing

DNA sequencing was performed by Macrogen (Seoul, Korea) using purified plasmid DNA samples and sequencing primers.

2.3 Protein expression

2.3.1 Transformation of plasmid

Vectors containing the autophagy-related genes were transformed into competent cells (Section 2.2.5) using the heat shock technique. Plasmid DNA (2 μL of 100 ng/μL stock) was added to a thawed 50-μL aliquot of competent cells and incubated on ice for 30 min. The cells were heat shocked in a digital heat block at 42°C for 35 s, after which they were immediately transferred to ice. SOC medium (250 μL) was added, and the mixture was incubated at 37°C for 60 min with shaking at 140 rpm. A 50-μL aliquot of cells was spread on a pre-warmed LB agar plate

containing appropriate antibiotics and incubated at 37°C for 12–16 h. Resulting colonies were used to prepare glycerol stocks or for subsequent plasmid preparation.

2.3.2 Small-scale expression

Small-scale expression trials were performed to determine the optimum incubation time and temperature for protein expression. Two methods of induction were trialled: isopropyl β-D-1-thiogalactopyranoside (IPTG) and auto-induction (Baneyx, 1999).

2.3.2.1 Bacterial expression

2.3.2.1.1 IPTG-induced protein expression

Colonies from LB agar plates were used to inoculate LB medium (10 mL) containing appropriate antibiotics. Suspension cultures were grown in centrifuge tubes (50 mL) to an OD_{600} of 0.4–0.6 before IPTG was added (final concentration 0.2 mM or 1 mM). To assess the effect of temperature on protein expression, the cultures were incubated at either 26°C or 37°C, with shaking at 180 rpm. Samples (500 µL) were taken at various time points, centrifuged (12,000 rpm, 5 min), and the supernatant was removed. SDS-PAGE buffer (80 µL) was added to the pellets, which were subsequently boiled for 5 min to lyse the cells, and 10 µL were loaded onto an SDS-PAGE gel.

2.3.2.1.2 Protein expression by auto-induction

Similar to the IPTG induction protocol, colonies from LB agar plates were used to inoculate auto-induction medium (10 mL) containing antibiotics, and incubated at 26°C and 37°C with shaking at 180 rpm. Samples were collected as described above.

2.3.2.2 Mammalian expression

Small-scale mammalian expression trials were performed using the FreeStyle™ 293 Expression System (Thermo Fisher Scientific) following the manufacturer's protocol. FreeStyle™ 293-F cells, derived from human embryonic kidney (HEK) cells, were thawed in a 37°C water bath and transferred into a disposable 125-mL polycarbonate Erlenmeyer shaker flask containing 17 mL of pre-warmed FreeStyle™ 293 Expression medium. The cells were incubated at 37°C in a humidified atmosphere of 8% CO_2 in air. The cap of the flask was loosened to allow aeration, and the flask was shaken at 125 rpm. The cells were subsequently subcultured until transfection.

Lipid-DNA transfection complexes were prepared by adding 30 μg of plasmid DNA to Opti-MEM®I medium in a total volume of 1 mL, and by adding 60 μL of 293fectin™ transfection reagent to 940 μL of Opti-MEM®I, incubating both solutions for 5 min, then mixing them together. The mixture was incubated for 30 min at room temperature to allow the formation of DNA-293fectin™ complexes.

Cells were diluted into a fresh flask with FreeStyle™ 293 Expression medium to obtain a suspension containing 3×10^7 cells in a volume of 28 mL. The DNA-293fectin™ complex was added to the flask to obtain a total volume of 30 mL, with approximately 1×10^6 viable cells/mL. The cells were incubated as above until they were harvested approximately 48 h post-transfection. Protein expression was checked by SDS-PAGE (Section 2.5.3).

2.3.3 Large-scale expression

For large-scale production of recombinant proteins, BL21 (DE3) cells containing the vector with gene of interest were streaked on LB agar plates containing

appropriate antibiotics and incubated overnight at 37°C. A single colony was used to inoculate 10 mL of LB broth (with antibiotics) and incubated overnight (37°C, 180 rpm) as a starter culture.

The starter culture was used to inoculate 200–400 mL of LB medium in a 2-L baffled flask. The culture was incubated at 37°C with shaking at 180 rpm until the OD_{600} reached 0.4–0.6, and then protein expression was induced with 0.2 mM IPTG. The culture was then incubated overnight at 26°C with shaking at 180 rpm. Cells were harvested by centrifugation at 4000 rpm for 10 min, and then resuspended in 20 mL of buffer B. The cell suspension was either used to purify the recombinant protein or snap frozen in liquid nitrogen and stored at −80°C.

2.4 Protein purification

2.4.1 Protein extraction

2.4.1.1 Sonication

Sonication uses a small metal probe oscillating at ultrasonic frequency to generate localised low pressure and membrane disruption through cavitation. Cells were lysed using a Sonicator 3000 (Misonix, Farmingdale, NY, USA) equipped with a microprobe (Misonix) for volumes less than 5 mL, and a standard probe for larger volumes (30 mL). Cell suspensions were prepared in a 50-mL plastic beaker and placed in an ice slurry to prevent excessive heating of the sample. The cells were sonicated at 60% amplitude with a pulse interval of 3 s on and 10 s off for 20 min.

2.4.1.2 High pressure homogenisation

Cells can be lysed by homogenisation by passing a sample through a needle

valve at high pressure, causing cell disruption by shear stresses and decompression as cells return to normal atmospheric pressure. Volumes larger than 30 mL were lysed by passing the sample through a Microfluidics M-110P cell disruptor (Newton, MA, USA) at 17,000 psi and 4°C. The cycle was repeated three times to ensure complete lysis.

The cell debris from the lysed samples was cleared by centrifugation at 15,000 rpm for 25 min at 4°C. The supernatant, containing the soluble recombinant protein, was then purified by chromatography steps, as described in the following sections.

2.4.2 IMAC for His_6-tagged proteins

IMAC was used to purify His_6-tagged proteins. Both Ni-NTA sepharose resins (GE Healthcare) and pre-packed columns (Section 2.2.4) were used. The Ni-NTA resin was prepared by centrifugation at 600 rpm for 2 min. The storage buffer was removed by decanting the supernatant, and the resin was equilibrated by resuspending the pelleted resin in lysis buffer. To ensure proper equilibration, this centrifugation and resuspension process was repeated three times. The clarified cell lysate (Section 2.4.1) was filtered through a 0.22-μm syringe filter and mixed with 2 mL of Ni-NTA resin for 1 h at 4°C. The slurry was gently transferred to a PD-10 gravity column (GE Healthcare), and the unbound lysate was allowed to flow through the column. The resin containing bound protein was washed with 10 bed volumes (BV) of lysis buffer containing 1% Triton X-100. To remove all the unbound protein from the resin, the resin was washed with 10 BV of lysis buffer containing 1 M NaCl. Finally, the resin was washed with 20 BV of lysis buffer to remove the high-salt buffer. The resin was then resuspended with three BV of elution buffer, and the flow through containing the purified protein was collected. To elute protein without the His_6-tag, the resin was resuspended in 1 BV of lysis buffer, and 50 μL of 1 U/μL thrombin was added. The

mixture was incubated overnight at room temperature, then the purified protein without a His6-tag was collected.

For purification using a pre-packed column, the clarified lysate was filtered through a 0.22-μm syringe filter and applied to a 1-mL Hi-Trap Chelating column pre-equilibrated in buffer B (10 column volumes (CV)) using a Gilson peristaltic pump. Where appropriate, a HisTrap column (GE Healthcare) was used with the same protocol. The column was washed with 10 CV of lysis buffer to remove any unbound protein. Purified protein was eluted with a gradient set to 100% elution buffer on an ÄKTApurifier system (GE Healthcare), and the eluted fractions were stored at 4°C or snap frozen in liquid nitrogen and stored at −80°C for later use. All flow rates were maintained at 1 mL/min. Different transition metals were studied to determine the maximum affinity of the proteins for purification. The Hi-Trap Chelating column was charged with either 100 mM $CoSO_4$, 100 mM $ZnSO_4$, 100 mM $NiSO_4$ or 100 mM $CuSO_4$ for each trial, and the sample fractions were assessed by SDS-PAGE.

2.4.3 Purification using glutathione Sepharose for glutathione-S-transferase (GST)-tagged proteins.

Batch purification was used to purify the GST-tagged proteins. The binding of GST-tagged proteins depends on the size, conformation and concentration of the protein in the sample being loaded. Binding of GST to glutathione is also flow-dependent, and lower flow rates often increase binding capacity. Therefore, the increased incubation time in a batch purification mode increases the binding of the recombinant protein to the matrix for efficient separation from other soluble proteins in the cell lysate.

Glutathione Sepharose 4B (GE Healthcare) was prepared by centrifugation at 600 rpm for 2 min, and the storage buffer was removed by decanting the supernatant. The resin was equilibrated by resuspension in PBS buffer, then re-pelleting by centrifugation, as above, an additional three times. Approximately 2 mL of the prepared slurry was incubated with soluble lysate, extracted from 400 mL of bacterial culture (Section 2.5.1), for 90 min at room temperature with gentle shaking. The slurry was then centrifuged at 600 rpm for 5 min to pellet the resin with bound protein. The supernatant was discarded, and the pelleted resin was gently transferred to a PD-10 gravity column (GE Healthcare). Unbound protein was removed by washing with PBS + 1% Triton X-100 and with PBS + 1 M NaCl (five CV of each). The slurry was equilibrated with PBS before adding 80 μL of thrombin (1 U) and incubating for 16 h at room temperature. The cleaved protein was eluted in PBS (three CV). Generally, the eluted proteins were further purified by preparative SEC (Section 2.4.4).

2.4.4 SEC

SEC was performed as an additional purification step to separate proteins by size to remove unwanted proteins that co-eluted with target protein in the initial IMAC or GST purification steps. A prepacked 24-mL Superdex 200 gel filtration column (GE Healthcare Life Sciences) was connected to an AKTAxpress chromatography system (GE Healthcare Life Sciences) and washed with three CV of MQ water before being equilibrated with 75 mL of SEC buffer at a flow rate of 0.3 mL/min. The concentrated protein sample was then loaded onto the column and run at a flow rate of 0.25 mL/min at 4°C for a total buffer volume of 30 mL, then 1-mL fractions were collected in a 96-well plate. Fractions that corresponded to the A_{280} peak on the chromatograph (UniCorn 5.11 Workstation, GE Healthcare Life

Sciences) were analysed by SDS-PAGE (Section 2.5.3), and fractions containing pure target protein were pooled and concentrated using a Vivaspin® 6 10,000 Da molecular weight cut-off (MWCO) polyethersulfone spin concentrator (Sartorius, Goettingen, Germany). Pure protein was snap frozen in liquid nitrogen and stored at −80°C.

2.5 Protein analysis

2.5.1 Protein concentration

Purified protein concentration was determined by measuring absorption at 280 nm (A_{280}) on a NanoDrop 2000 spectrophotometer. Extinction coefficients of the specific proteins were calculated from their amino acid sequence using the ExPasy server (Gasteiger *et al.*, 2005), and were used to determine protein concentration using the equation:

$$\text{Concentration} = (A_{280}/\varepsilon_{280}) \times l$$

Where A_{280} is the absorption at 280 nm in AU, ε_{280} is the extinction coefficient at 280 nm ($M^{-1}.cm^{-1}$), and l is the path length.

2.5.2 Buffer exchange

(i) Buffer exchange by dialysis

Protein samples were transferred into dialysis tubing (10,000 MWCO), and dialysed against the required buffer. Sample to dialysis buffer ratios were typically 1:2000. Dialysis was performed at 4°C with stirring to increase the rate of exchange. Typically, this was repeated to ensure complete exchange.

(ii) Buffer exchange by chromatography

HiTrap Desalting (5 mL) columns (GE Healthcare) were pre-packed with Sephadex G-25 Superfine cross-linked dextran. These employ the same principles as gel filtration to separate molecules with a mass larger than 1 kDa from those with a mass below 5 kDa, thereby separating proteins from buffer salts. The column was connected to an ÄKTApurifier system, and equilibrated with three CV of the required buffer at 2 mL/min. A maximum of 1 mL of sample was loaded onto the column. Protein was eluted at 2 mL/min using the required buffer, 0.5 mL fractions were collected and protein elution was monitored at A_{280}.

(iii) Buffer exchange by dialysis devices

Slide-A-Lyzer dialysis devices (Thermo Fisher Scientific) with a 7,000 MWCO were used for faster buffer exchange of smaller volumes of sample (<1 mL). Protein sample (100–500 µL) was pipetted into the dialysis cup and partially submerged in the appropriate buffer. Dialysis was performed at room temperature or 4°C, with stirring to increase the rate of exchange.

2.5.3 SDS-PAGE

SDS-PAGE is the most common technique to check the purity of protein samples. It completely denatures proteins in the presence of SDS, an ionic detergent, which breaks hydrogen bonds and coats the entire protein to give it a negative charge. This means protein migrates through the gel in a molecular weight-dependent manner. In the presence of reducing agents, such as dithiothreitol or β-mercaptopethanol, any disulfide bonds are reduced, thus denaturing the protein molecule.

NuPAGE® 4–12% Bis-Tris gels were run according to the manufacturer's recommendations in a NuPAGE® gel electrophoresis box at room temperature. Novex® Sharp protein ladder (3 µL; Life Technologies) was used as a molecular

weight marker. Protein samples were mixed with reducing agent and 4× lithium dodecyl sulphate (LDS) sample buffer or 2× Tris-Gly sample buffer, depending on the nature of the gel being run. Protein samples were typically denatured at 70–90°C for 5 min in SDS-PAGE buffer before loading onto the gel. After electrophoresis, gels were placed in distilled water and microwaved for 20 s to remove the SDS present in the running buffer. They were then placed in Simply Blue stain for 1 h, or until protein bands appeared, and subsequently destained in distilled water for 1 h.

Table 2.13: Solutions used for SDS-PAGE.

Solution	Content
MES running buffer	50 mM Tris-MES (pH 7.6) 0.1% SDS, 1 mM EDTA.
LDS loading dye	564 mM Tris base, 424 mM Tris-HCl, 40% glycerol, 8% LDS, 2.04 mM EDTA, 0.88 mM Coomassie Blue G-250, 0.7 mM phenol red (pH 8.5).
Reducing agent	1 M β-mercaptoethanol.
Tris-Gly loading dye	126 mM Tris-HCl (pH 6.8), 20% glycerol, 0.005% bromophenol blue.
Simply Blue stain	80 mg/L Coomassie Brilliant Blue in 35 mM HCl.

2.5.4 Differential scanning fluorimetry (DSF)

The thermal stability of a protein can be determined using DSF. Pure protein samples were incubated with the fluorescent dye SYPRO® Orange (Invitrogen) in clear, thin-walled 96-well PCR plates. The samples were then heated using an IQ™ 5 Real-time PCR Detection System (Bio-Rad), which measures fluorescence. Heating of the protein sample unfolds it and subsequently exposes its hydrophobic regions, which then bind the fluorescent dye, and this increase in fluorescence is measured as the temperature increases. SYPRO® Orange was excited at 490 nm, and its emission was measured at 595 nm. A stock solution containing 1 μL of SYPRO® Orange, with protein and buffer added up to a final volume of 100 μL, was prepared. Stock solution (25 μL) was added to each well, and assays were performed in triplicate. Themostability curves were obtained by heating the samples from 20°C to 100°C in

0.5°C increments, with a 10-s dwell time at each temperature. Data was processed and analysed using IQ™ 5 software (Bio-Rad).

2.5.5 SEC-multi-angle light scattering

Static light scattering (SLS) is a technique for determining the absolute molecular mass of a protein sample in solution through exposure to a low intensity laser. The intensity of the scattered light is measured as a function of angle, and the data are analysed to calculate the molecular mass. In addition to its use for quality control of proteins, SLS is a precise method for determining the oligomeric state of a protein. Because light scattering and concentration are measured for each of the eluting fractions, the mass and size can be determined independent of the elution position. This is particularly important for novel protein architectures that may elute at positions distant from those predicted by the calibration for the SEC column.

The weight-average molecular weight (Wen *et al.*, 1996) is given by:

$$M = K' \frac{(LS)}{(RI)}$$

Where K' is given by:

$$K' = \frac{K_{RI}}{K_{LS}(dn|dc)}$$

K_{RI} and K_{LS} are instrument calibration constants, LS is the light scattering signal, RI is the refractive index signal and dn|dc is the refractive index increment, which is defined as 0.186 mL/g for globular proteins (Wen *et al.*, 1996).

A Superdex 200 GL 10/300 column (GE Healthcare) was connected to a Viscotek 302-040 Triple Detector GPC/SEC system (Malvern Instruments Ltd., Malvern, UK), operated at 28°C, and equilibrated with appropriate buffer at 0.4

mL/min for a minimum of three CV until a stable baseline was observed. Protein samples (110 µL) were loaded onto the column. Absolute molecular weight, radius of hydration and size distributions were calculated using the refractive index (RI), intrinsic viscosity, and right-angle light scattering (RALS) measurements calibrated against bovine serum albumin (66.5 kDa; Sigma-Aldrich), which was run at the beginning and end of each sample sequence. Analysis was performed using OmniSEC software (Malvern Instruments) as per the manufacturer's instructions.

2.5.6 Pull-down assay

Pull-down assays can verify the interaction between two proteins. It is a simple affinity purification technique in which a "bait" protein, expressed with a tag, is immobilised to beads and incubated with a solution containing the "prey" protein.

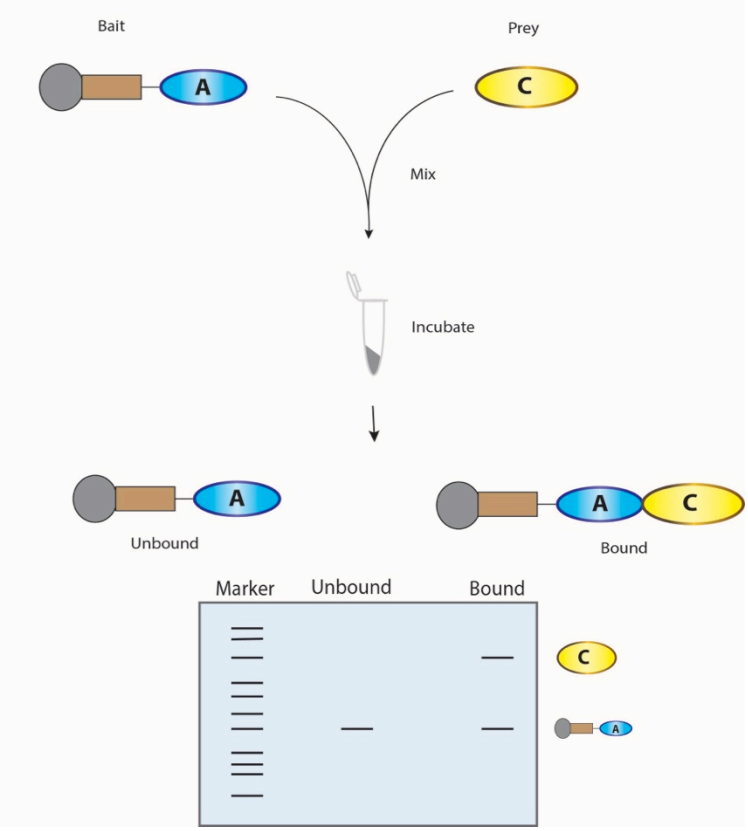


Figure 2.1: Graphical representation of the pull-down assay procedure where the bait (blue) protein, expressed with a tag and bound to the resin, captures prey (yellow) protein. The interaction is verified using SDS-PAGE.

In this study, the bait protein was expressed with a GST-tag and was bound to glutathione Sepharose 4B resin (GE Healthcare). Glutathione Sepharose 4B resin (20 μL) was added to a 1.5-mL microcentrifuge tube and washed twice with 500 μL of PBS buffer by centrifuging and discarding the supernatant. Prey protein (100 μL) was added to the washed glutathione Sepharose 4B resin, and incubated for 1 h at 4°C. Following incubation, the beads were washed twice with 500 μL of PBS + 0.1% Triton X-100 buffer, then the resin was washed twice with PBS (without Triton X-100). SDS-PAGE loading dye was added to the resins, and the results were analysed on an SDS-PAGE gel.

2.5.7 X-ray crystallography

2.5.7.1 Crystallisation

The PLEKHM1-LC3A, PLEKHM1-GABARAP and PLEKHM1-GABARAPL1 chimeric proteins were purified and crystallised as chimeric LIR-fusion proteins, whereas the $LC3C^{8–125}$ region was co-crystallised with the PLEKHM1 LIR peptide (629–PQQEDEWVNVQYPD–642). Initial crystallisation trials were performed using the following crystallisation screens: Crystal Screen, Crystal Screen Cryo, Index, PEG/Ion (all acquired from Hampton Research, Alisa Viejo, CA, USA), JCSG+, Midas, Morpheus, PACT, Clear Screen Strategy I and Clear Screen Strategy II(all acquired from Molecular Dimensions, Suffolk, UK). In all cases, the drops included 400 nL of protein (concentrations listed below) and 400 nL of mother liquor. All crystallisation experiments were set up and stored at 4°C.

PLEKHM1-LC3A (10 mg/mL) crystals were grown in JCSG+ screen condition H7 (0.2 M ammonium acetate, 0.1 M Bis-Tris (pH 5.5), 25% w/v polyethlene glycol 3,350). Crystals for PLEKHM1-GABARAP (9.1 mg/mL) were

grown in PEG/Ion screen condition F5 (4% v/v Tacsimate (pH 8.0), 12% w/v polyethylene glycol 3,350). Crystals for the PLEKHM1-GABARAPL1 protein (7.5 mg/mL) were grown in PEG/Ion screen condition A6 (20% w/v polyethylene glycol 3,350, 0.2 M NaCl, 8% 2-methyl-2,4-pentanediol (pH 7.2)). The $LC3C^{8–125}$ protein (9.2 mg/mL) was mixed with the PLEKHM1 peptide (2.4 mg/mL) in equal volumes and incubated for 3 h at 4°C, prior to setting up the crystallisation trays. Crystals were grown in PEG/Ion screen condition D5 (0.2 M potassium phosphate monobasic, 20% w/v polyethylene glycol 3,350). The crystals were frozen in liquid N_2 prior to data collection.

The ULK1-GABARAP chimeric protein was purified and crystallised as a chimeric LIR fusion protein. Initial crystallisation trials were performed using the following screens: Crystal Screen, Crystal Screen Cryo, Index and PEG/Ion (Hampton Research), JCSG+, Midas, Morpheus, PACT, Clear Screen Strategy 1 and Clear Screen Strategy 1 (Molecular Dimensions). The drops included 400 nL of protein (concentration listed below) and 400 nL of mother liquor. All crystallisation experiments were set up at 4°C and room temperature. The ULK1-GABARAP (10 mg/mL) crystals were grown in Crystal Screen Cryo Screen condition G6 at room temperature (0.08 M HEPES (pH 7.5), 8% w/v polyethylene glycol 6,000, 4% v/v (+/−)-2-ethyl-2,4-pentanediol, 20% v/v glycerol).

2.5.7.2 Data collection and processing

X-ray diffraction data was collected on the MX2 micro-crystallography beamline at the Australian Synchrotron (Melbourne, Australia). The data were integrated using XDS (Kabsch, 2010a, b) and scaled using Aimless (Evans & Murshudov, 2013). The structures were solved by molecular replacement using MOLREP (Vagin & Teplyakov, 1997), and phases for the co-crystal structure were

estimated using PHASER (McCoy *et al.*, 2007). The solved structures were refined using PHENIX.REFINE (Adams *et al.*, 2010), and model building was performed using COOT (Emsley *et al.*, 2010). Images were generated using PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.).

2.5.8 Surface plasmon resonance (SPR)

All SPR experiments were performed on a ProteOn™ XPR36 Protein Interaction Array System using ProteOn™ GLC Sensor Chips at 25°C and ProteOn™ 0.35-mL or 2-mL deep-well 96-well plates (all purchased from Bio-Rad). Following optimisation of ligand immobilisation levels, buffer conditions, analyte concentrations, injection parameters and regeneration conditions, SPR experiments were performed as described below.

2.5.8.1 Buffer and protein solution preparation

All buffers and solutions used for SPR were prepared fresh, and then filtered using 0.2-µm syringe-driven or vacuum-driven filters, before being sonicated for 20 min (Digitech Ultrasonic Cleaner, Digitech Systems, Kolkata, India) to degas and disperse any remaining particulate in solution. Activation and deactivation solutions for amine coupling were supplied as part of the ProteOn™ amine coupling kit (Bio-Rad). All protein samples were centrifuged at 4°C for 5 min at 12,000 rpm to pellet any aggregate that may have been present in the sample, prior to the concentration being checked in triplicate as described in Section 2.5.1. PBS with 0.05% Tween-20 (PBST) was used as the running buffer in all experiments and in the blank buffer injections.

2.5.8.2 Sensor chip initialisation and pre-conditioning

Each time a new or used sensor chip was inserted into the SPR machine, a wet initialisation step (50% v/v glycerol) was performed. Following several system flushes with running buffer, the sensor chip surface of a new chip was conditioned prior to any immobilisations with successive 30-µL pulses of 0.25% Tween-20, 50 mM NaOH and 100 mM HCl for 60 s each at a flow rate of 30 µL/min, first in the vertical direction (ligand), followed by the horizontal direction (analyte). Blank buffer runs (eight in the analyte direction then eight in the ligand direction) were performed to remove all traces of the pre-conditioning solutions and to equilibrate the chip surface (100 µL of PBST at a flow rate of 100 µL/min for 60 s).

2.5.8.3 Interaction of the ULK1 LIR with LC3A and GABARAP wild type and mutant proteins

Two approaches were used to investigate the interaction between LC3A, LC3A-R70L, LC3A-F52Y, GABARAP, GABARAP-R67L and GABARAP-Y49F with the ULK1 LIR. In the first approach, an anti-GST polyclonal antibody was immobilised on the chip surface in all six ligand directions, and the recombinant GST-tagged ULK1 LIR protein was allowed to interact with the antibody. The LC3A and GABARAP proteins were injected in the analyte direction to measure the interaction of the ULK1 LIR with LC3A and GABARAP.

In the second approach, the ULK1 LIR peptide (KSCDTDDFVMVPAQ) and a scrambled ULK1 LIR control peptide (KVSFPQACDMDTVD) were immobilised on the surface of a GLC chip (Bio-Rad) in channels L2 and L5 (control), and L3 and L6 (ULK1 LIR), by amine coupling (Fischer, 2010). After optimising peptide concentrations and activation and deactivation strategies to ensure a meaningful response from the interaction, the ligand was immobilised on the chip surface. The

LC3A and GABARAP proteins were injected in the analyte direction to measure the interaction of the ULK1 LIR with LC3A and GABARAP.

2.5.8.3.1 Immobilisation

The carboxymethyl dextran on the sensor chip surface was activated to produce reactive succinimide esters on the ligand channels of the sensor chip surface by injecting 150 µL of a 1:1 mixture of 0.4 M 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide and 0.1 M N-hydroxysuccinimide (NHS) at a flow rate of 30 µL/min for 300 s. Ligand was then immobilised on each channel by injecting 150 µL of 0.75 µg/mL antibody (in 10 mM sodium acetate buffer (pH 4)) for 300 s at a flow rate of 30 µL/min. Ligand channel 1 (L1, activation and deactivation only) was used as a reference channel, and ligand channel 4 (L4, no activation or deactivation) was used as a blank channel. The ULK1 LIR peptide was immobilised in L3 and L6, while the control peptide was immobilised in L2 and L5. Ligand immobilisation was directly followed by deactivation of unreacted succinimide esters with 1 M ethanolamine HCl (pH 8) at a flow rate of 30 µL/min for 300 s.

Following two blank buffer injections, four stabilisation steps were performed in the ligand direction to stabilise the ligand surface and remove any weakly bound ligand molecules. Two 30-µL pulses of 10 mM sodium acetate (pH 3.0) were followed by two 30-µL pulses of 10 mM glycine HCl (pH 2.0), with each 18-s injection having a flow rate of 100 µL/min. Multiple blank buffer injections, first in the ligand direction and then the analyte direction, were performed until a stable baseline was achieved across all channels.

2.5.8.3.2 Interaction of LC3A wild type, LC3A-R70L and LC3A-F52Y with the ULK1 LIR

LC3A-wt, LC3A-R70L and LC3A-F52Y were injected in the analyte direction at six different concentrations: 25 μM, 12.5 μM, 6.25 μM, 3.125 μM, 1.56 μM and 0.78 μM. A 275-μL volume of each concentration was injected across the chip at a flow rate of 60 μL/min and allowed to associate with the ULK1 LIR, followed by dissociation for 600 s and flow stabilisation for 20 s. The ligand surface was then regenerated with a 30-μL pulse of glycine HCl (pH 1.0) for 18 s at a flow rate of 100 μL/min to remove any undissociated analyte. At least four blank buffer injections were performed to allow the ligand to recover and re-establish a stable baseline. The interaction for each analyte was repeated three times.

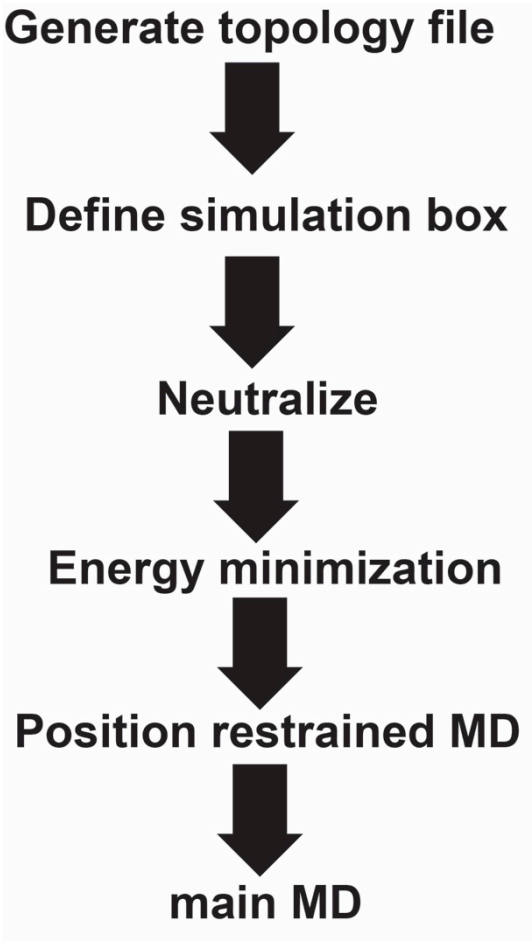
2.5.8.3.3 Interaction of GABARAP wildtype, GABARAP-R67L and GABARAP-Y49F with the ULK1 LIR

GABARAP-wt, GABARAP-R67L and GABARAP-Y49F were injected in the analyte direction at six different concentrations: 1.56 μM, 0.78 μM, 0.39 μM, 0.195 μM, 0.097 μM and 0.048 μM. Experimental conditions were as described in Section 2.5.8.3.2. The data was processed using Scrubber and ClampXP software to obtain the kinetics and binding affinity.

2.5.9 Molecular dynamics simulation

Molecular dynamics simulations were performed using GROMACS (version 4.6.3) (Berendsen *et al.*, 1995, Lindahl *et al.*, 2001, Van Der Spoel *et al.*, 2005, Hess *et al.*, 2008). Simulations were performed with coordinates extracted from the structures of LC3A (PDB: 3WAL), GABARAP (PDB: 1GNU), the Atg13 LIR fused to LC3A (PDB: 3WAN) and the ULK1 LIR fused to GABARAP (solved in this

study). The Atg13 LIR amino acid sequence (FVMI) was modified to the ULK1 LIR sequence (FVMV) using COOT model building software (Emsley *et al.*, 2010). Simulations were performed using the workflow below:



The topology file was generated using the pdb2gmx command with the coordinate file as the input. The topology file contains information about the atoms, such as bond length and bond angles. The force field GROMOS96 54a7 (Schmid *et al.*, 2011) was chosen because it is a general-purpose simulation package for biomolecules, and the simple point charge water model was selected during the generation of the topology file, as recommended by the GROMACS instructions. Next, the simulation box was defined using the editconf, genbox and grompp commands, where "cubic" was selected as the box type and the molecule was centred in the box. The distance between the molecule and the box was assigned as 0.9 nm.

The system was set up in a water environment neutralised with 0.15 M NaCl using the genion command.

Addition of solvents and hydrogens causes interference with the structure. Therefore, energy minimisation was performed to bring the system to a relaxed state. The system was prepared for energy minimisation using grompp, followed by an mdrun of 50,000 energy-minimisation steps. Energy minimisation was performed until a maximum force of less than 1000 kJ/mol/nm was achieved. A group cut-off of 1.4 nm was implemented for the van der Waals (vdw) settings.

Position-restrained molecular dynamics (MD) was performed after energy minimisation. During this process, the atom positions are restrained and the solvent is allowed to move freely through the structure and create a proper environment such that the structure is properly immersed (or soaked) in the water. Position-restrained MD was performed for 1,500,000 steps (3 ns) using a group vdw cut-off of 1.4 nm. V-rescale thermostat was used to perform the simulation at fixed temperature.

Finally, the main MD simulations were calculated for 100 ns (50,000,000 steps) at 300 K using computer clusters. After the simulation, the molecule was placed at the centre of the box using the trjconv command. The root mean square deviation (RMSD) and the root mean square fluctuation (RMSF) were calculated using the trajectory file obtained from the simulation as the input file, with the gmx rms and gmx rmsf commands, respectively. The bond distances between the atoms over 5000 frames were calculated using Visual Molecular Dynamics (VMD) (Humphrey *et al.*, 1996). The frame simulations were visualised using VMD, and important frames that displayed interesting interactions were saved as pdb files. Images from the frames saved in pdb format were prepared using PyMOL.

Chapter 3. CLONING, EXPRESSION, PURIFICATION AND INTERACTION STUDIES ON FIP200 AND ATG13

3.1 Introduction

The aim of this thesis was to study the interactions between the proteins involved in the autophagic pathway. In this chapter, the region of autophagy-related protein 13 (Atg13) that interacts with focal adhesion kinase family interacting protein of 200 kDa (FIP200) was studied. Atg13 and FIP200 are part of the human autophagy initiation complex, called the UNC51-like Kinase 1 (ULK1) complex (Hara *et al.*, 2008, Hosokawa *et al.*, 2009). During starvation, inhibition of mammalian target of rapamycin (mTOR) dephosphorylates ULK1 and Atg13, which activates the ULK1 kinase activity to phosphorylate FIP200 (Metcalf *et al.*, 2012). FIP200 interacts with the C-terminus of Atg13, but the region of FIP200 that interacts with Atg13 is unknown (Hosokawa *et al.*, 2009). To identify this, studies of the interaction between the C-terminus of Atg13 and fragments of FIP200 were attempted.

Unlike the yeast initiation complex (the Atg1 complex), the members of the ULK1 complex (ULK1, FIP200, Atg13 and Atg101) remain stably associated in both the active and inactive states (Hara *et al.*, 2008, Hosokawa *et al.*, 2009). Interestingly, Atg13 keeps the entire ULK1 complex stably associated by mediating the interaction between ULK1 and FIP200, and also mediates the phosphorylation of FIP200 (Hosokawa *et al.*, 2009, Jung Ch Fau - Jun *et al.*, 2009, Metcalf *et al.*, 2012). Therefore, studying the interaction between FIP200 and Atg13 provides a starting point for studying the full ULK1 complex, and is an initial step towards understanding the mechanism of ULK1 complex interaction.

In addition to their role in the initiation of starvation-induced autophagy, FIP200 and Atg13 are functionally important proteins. FIP200 is a multifunctional protein that is involved in various cellular activities, such as cell proliferation, cell migration, cell size regulation, cell death and tumour suppression (Chano *et al.*, 2002, Gan *et al.*, 2005, Gan *et al.*, 2006, Melkoumian *et al.*, 2005). Silencing of either FIP200 or Atg13 inhibits autophagosome formation, thereby making these proteins important targets for understanding diseases related to autophagy (Hara *et al.*, 2008, Hosokawa *et al.*, 2009, Klionsky *et al.*, 2007, Rubinsztein, 2006, Rubinsztein *et al.*, 2012, Rubinsztein *et al.*, 2011).

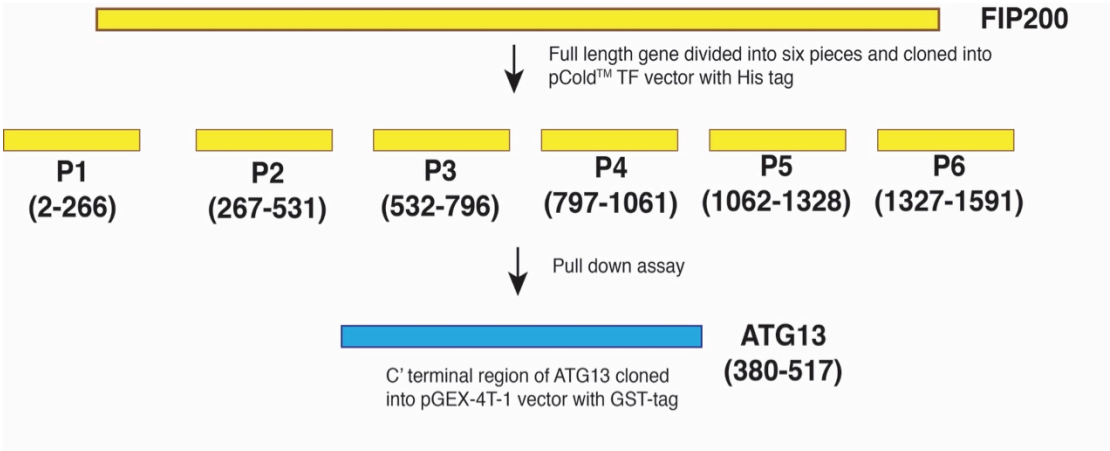


Figure 3.1: Schematic representation of the strategy used to study the interaction between FIP200 and Atg13. The FIP200 gene was truncated into six fragments (labelled P1–P6 and shown in yellow). Each of the six fragments was cloned into the pCold™ TF vector, which encodes an N-terminal His_6 tag, followed by the sequence of the chaperone trigger factor, in front of the gene of interest. The proteins expressed by these six constructs were used as "prey" in a pull-down assay. The C-terminal of Atg13 (380–517), shown in blue, was cloned into vector pGEX-4T-1 with a GST tag, and the expressed protein was used as "bait" in the pull-down assay.

In this chapter, the FIP200 gene was divided into six equal-sized truncated regions to map the Atg13-interacting region in FIP200 (Figure 3.1). This strategy was used because of the ease of identifying the binding region from a smaller sequence range, and because the relatively large size of the full-length protein (180 kDa) reduces the probability of successful expression in bacterial cells. The six FIP200 constructs were expressed with N-terminal His_6-trigger factor (TF) sequences. TF is a

ribosome-associated protein chaperone found in prokaryotes that promotes folding of newly-expressed polypeptides (source: Takara Bio Inc., Shiga, Japan). The C-terminal Atg13 region was expressed with a GST-tag (Figure 1.1). Despite performing a pull-down assay to identify the FIP200 construct that interacted with the C-terminal ATG13 construct, the C-terminal Atg13 construct did not bind any of the six FIP200 constructs.

3.2 Results

3.2.1 Cloning and expression of full-length Atg13

A high level of expression of correctly folded and stable Atg13 was required to study the interaction between Atg13 and FIP200, and to study the crystal structure of full-length Atg13. Full-length Atg13 (517 amino acids) is a challenging protein to express using a simple expression system like *E. coli* because it is a complex eukaryotic protein, with nine cysteine residues that require chaperones for proper folding, and undergoes heavy post-translational modification. Therefore, a mammalian expression system was chosen for expression of full-length Atg13.

The full-length human ATG13 gene was cloned into a pCMV-Fc fusion vector to express the ATG13 protein fused to an immunoglobulin Fc domain in Freestyle human embryonic kidney 293 (HEK293) cells. PCR amplification of the full-length human ATG13 gene generated a single 1557-bp DNA fragment (Figure 3.2A). Screening by colony PCR (Figure 3.2B), followed by DNA sequencing of the plasmid DNA from the colonies with correct-sized inserts, showed that a fragment with the correct sequence of the full-length ATG13 gene without any mutations had been cloned.

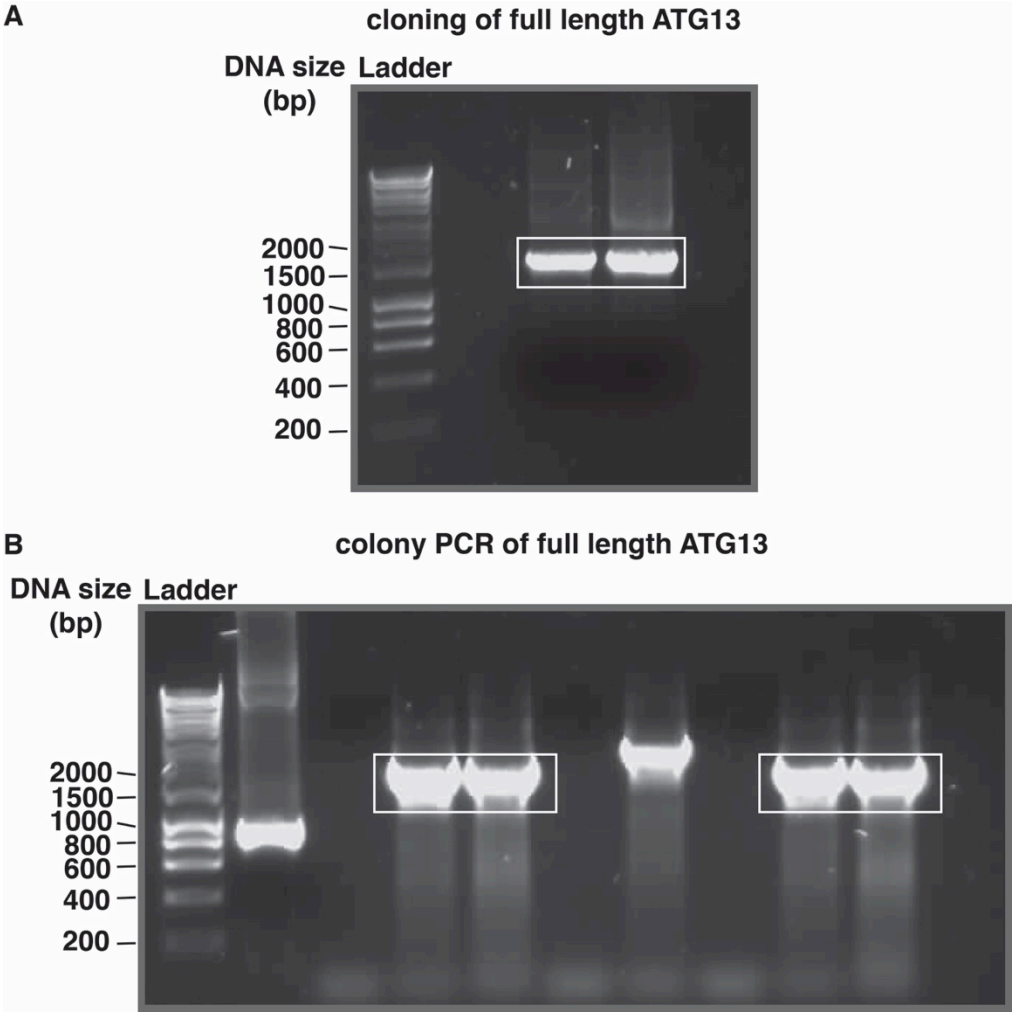


Figure 3.2: A. DNA agarose gel showing the amplified PCR products (gene encoding full-length ATG13) as distinct bands (marked with white box) that were cut from the gel for purification. Migration of a band at approximately 1500 bp suggests that the correct-sized fragment was amplified. B. Agarose gel showing colony PCR analysis, confirming that the gene encoding full-length ATG13 was successfully ligated into pGEX-4T-1 (observed as a band of ~1500–2000 bp). Plasmids from colonies with bands of the correct size were sequenced to verify ligation of the correct DNA sequence.

The expression of full-length ATG13 in HEK293 cells was unsuccessful. The pCMV-Fc fusion vector has a secretion signal that allows secretion of the expressed protein through endoplasmic reticulum/Golgi system into the culture medium. Human ATG13 is an intracellular protein, so there was a risk that the protein would be incompatible with the secretion and processing pathways, leading to its degradation. Therefore, removing the secretion signal and attempting to express full-length ATG13

as an intracellular protein may be successful. However, because of the time constraints and the high cost involved, this optimisation could not be performed in the current study.

3.2.2 Cloning, expression and purification of the C-terminal of ATG13 (380–517)

DNA encoding the region of ATG13 that binds FIP200 (amino acids 380–517) (Hosokawa *et al.*, 2009) was successfully cloned into pGEX-4T-1. PCR amplification generated a 411-bp product with BamHI and EcoRI restriction sites on either end. The amplified DNA fragment was observed as a discrete band on an agarose gel (Figure 3.3A).

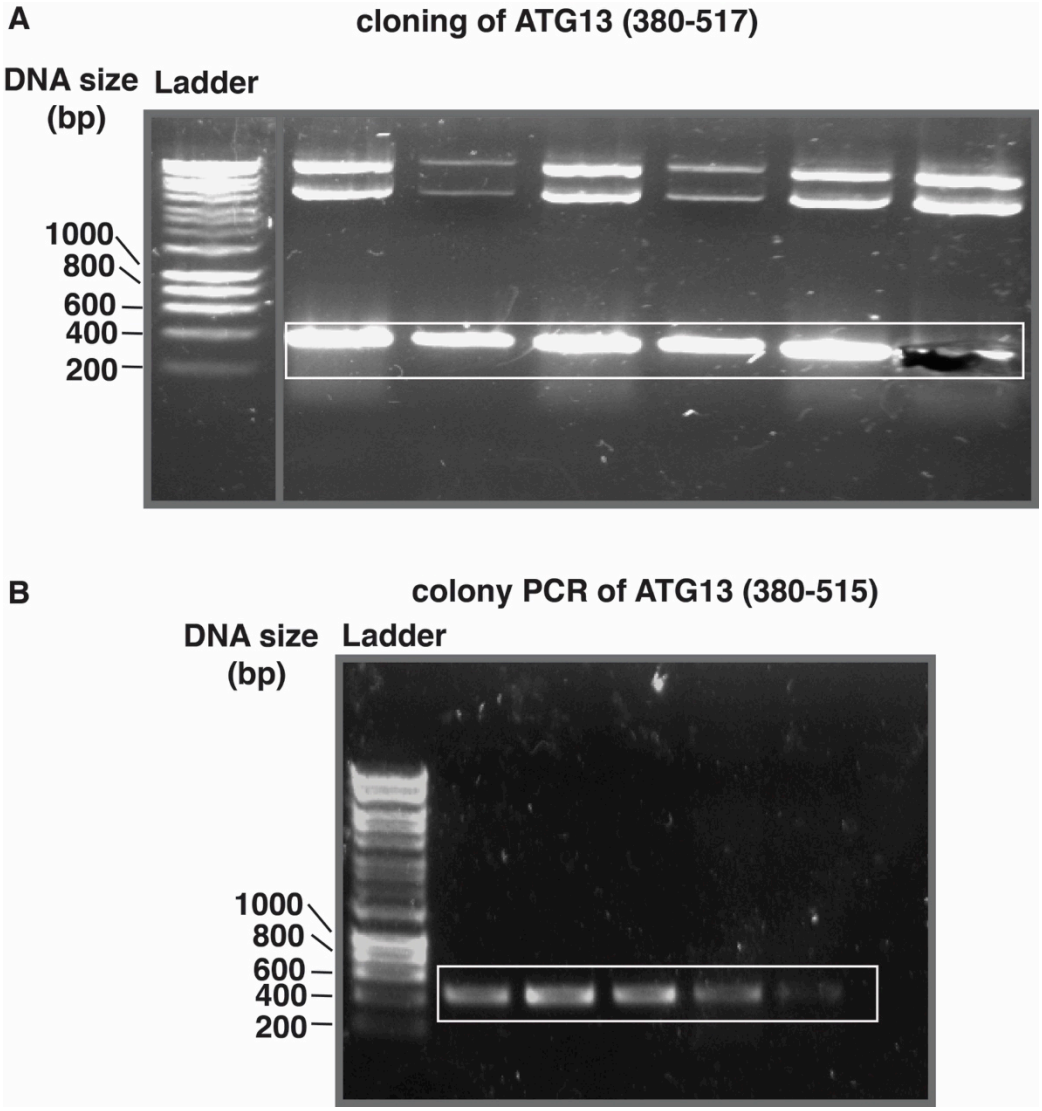


Figure 3.3: A. Agarose gel showing amplified PCR products (region encoding ATG13 residues 380–517) as distinct bands (marked with white box) that were cut from the gel for purification. No amplicon was obtained at the calculated annealing temperature, therefore temperature-gradient PCR was performed. Migration of a band at ~400 bp suggested that the correct-sized DNA fragment was amplified. B. Agarose gel analysis of colony PCR samples suggested that the region encoding ATG13 380–517 was successfully ligated into pGEX-4T-1, as demonstrated by a band of 400–600 bp (bands marked in white box). Plasmids from colonies producing the correct-sized bands were sequenced to verify ligation of the correct DNA sequence.

Colony PCR screening (Figure 3.3B) and subsequent sequencing of plasmids confirmed the presence of the correct sequence in the correct frame for expression. Therefore, the region of ATG13 that binds FIP200 was successfully cloned into vector pGEX-4T-1 for expression and interaction analysis.

Rosetta (DE3) cells, which contain the pRARE plasmid to supply tRNAs for rare codons, were used to achieve high-level expression of ATG13 amino acids 380–517. SDS-PAGE analysis of small-scale expression trials with the pGEX-4T-1 ATG13 expression vector showed a band at the expected size of 40 kDa. Despite this, the expression level was low compared with the FIP200 protein fragments (Figure 3.4A). Optimisation trials were performed to increase the expression level, with parameters including growth temperature, expression time, growth medium and IPTG concentration (Figure 3.4B), but these did not improve the expression level of the ATG13 (380–517) protein. Large-scale expression was performed with optimised conditions, and low levels of expression of the ATG13 (380–517) protein with a GST tag were obtained, despite trying various conditions (Figure 3.4C).

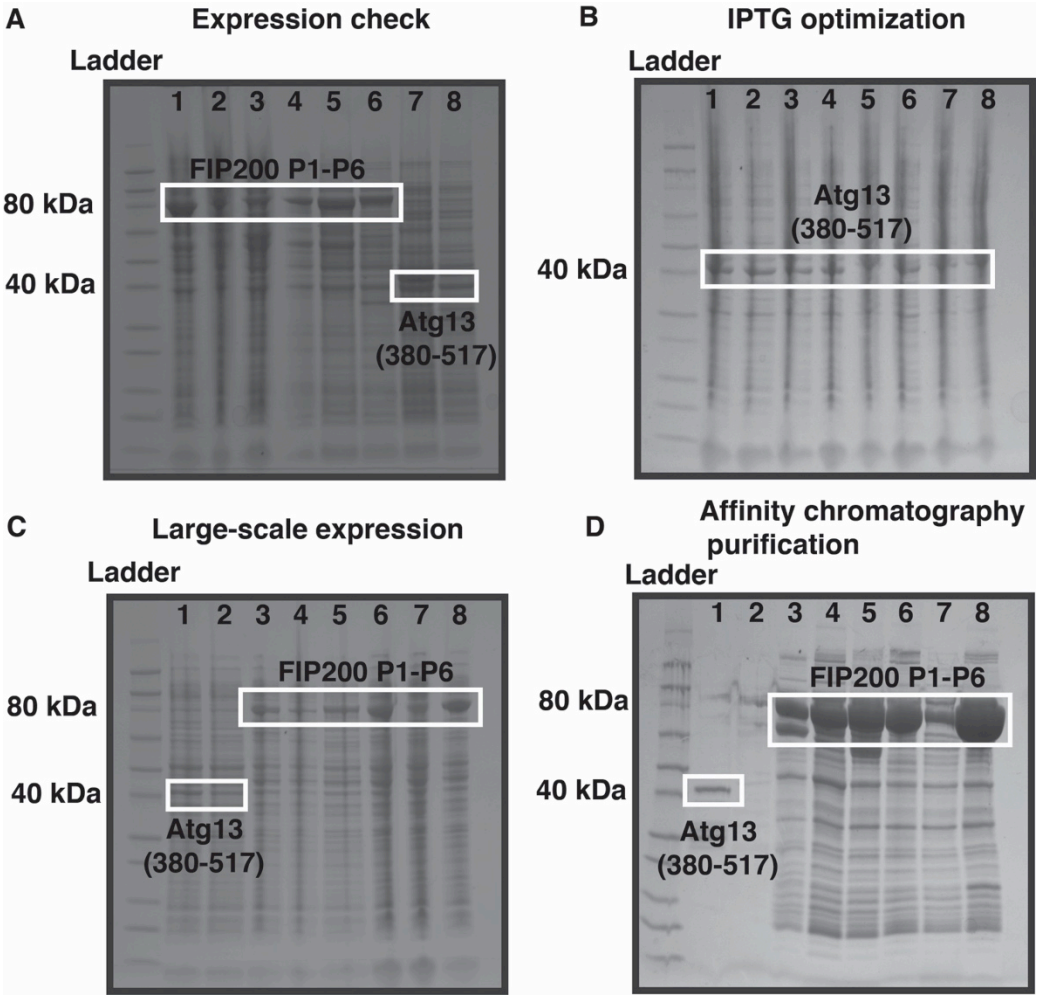


Figure 3.4: A. Samples from small-scale expression were analysed by SDS-PAGE, which demonstrated that the six FIP200 constructs and the ATG13 (380–517) construct were successfully expressed in *E. coli*. Lanes 1–6: crude cell lysate from BL21 (DE3) cells expressing FIP200 P1–P6, lane 7: cell lysate from Rosetta (DE3) cells expressing ATG13 (380–517), lane 8: cell lysate of BL21 (DE3) cells expressing ATG13 (380–517). B. To identify the ideal IPTG concentration for the induction of ATG13 (380–517), the cells were induced with different concentrations of IPTG. SDS-PAGE analysis showed that varying IPTG concentration did not affect the expression levels of ATG13 (380–517). Lanes 1–8: cell lysates from cultures induced with IPTG at the following concentrations: 0.1 mM, 0.2 mM, 0.3 mM, 0.4 mM, 0.5 mM, 0.6 mM, 0.8 mM and 1 mM. C. Samples from large-scale (1-L culture) expression were analysed by SDS-PAGE. SDS-PAGE analysis demonstrated that the six FIP200 constructs expressed well, whereas the ATG13 (380–517) construct showed low-level expression. Lanes 1 & 2: crude cell lysate from ATG13 (380–517), lanes 3–8: cell lysate from FIP200 P1–P6. D. SDS-PAGE gel showing samples from affinity chromatography purification. SDS-PAGE analysis showed that ATG13 protein expressed using Rosetta (DE3) was successfully purified, although the expression level was relatively low. All six FIP200 constructs were purified successfully. Lane 1: ATG13 (380–517) expressed using Rosetta (DE3) cells that were purified using glutathione Sepharose resin, lane 2: ATG13 (380–517) expressed using BL21 (DE3) cells that were purified using glutathione Sepharose resin, lanes 3-8: FIP200 (P1–P6) purified using Ni-NTA resins.

3.2.3 Strategy for cloning, expression and purification of FIP200 constructs

After careful analysis of secondary structures using XtalPred, the 4773-bp human FIP200 gene was truncated into six DNA fragments of around 800 bp each (exact size summarised in Table 2) to map the region of ATG13 that interacts with FIP200. The six FIP200 fragments expressed as polypeptides without a proper tertiary structure have the risk of becoming insoluble protein. Therefore, a chaperone-like TF tag was included to assist the folding of the polypeptide and obtain soluble protein. To achieve this, the six FIP200 DNA fragments were cloned into pCold™ TF, which is a cold-shock fusion vector that expresses the protein of interest as a TF (48 kDa) chaperone fusion. As TF is derived from *E. coli*, it is expressed highly in *E. coli* expression systems. The pCold™ TF vector also has a His_6 tag sequence before the TF sequence and a thrombin cleavage site, for potential TF removal, between the TF domain and the multiple cloning sites.

The amplified DNA fragments were verified on an agarose gel, which showed the 800-bp PCR product as a discrete band (Figure 3.5A and B). Colony PCR was performed to identify transformants containing plasmids with correct-sized inserts, and the resulting products were verified by agarose gel electrophoresis (Figure 3.6). Plasmids containing the correct insert produced a discrete 1000-bp band, including the multiple cloning site regions. Sequencing of the plasmid samples purified from the selected colonies confirmed the presence of the target gene, and showed that the target gene was in the correct frame for expression without any mutations in the sequence.

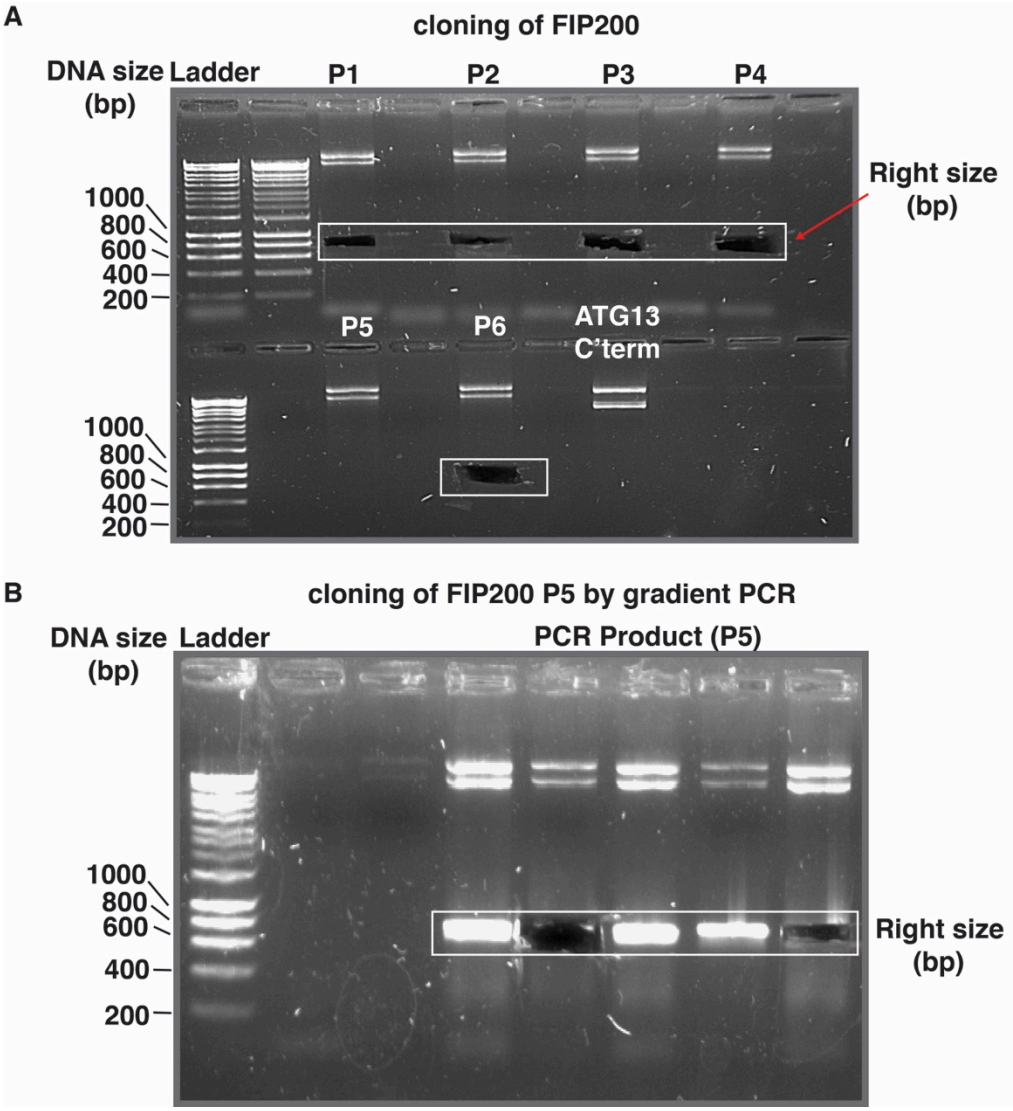


Figure 3.5: A. Typical agarose gel showing the amplified PCR products (FIP200 P1, P2, P3, P4 and P6) as distinct bands (marked with white box). The bands were cut from the gel to purify the amplified PCR product. The migration of the product to approximately 800 bp suggested that the correct DNA fragment was amplified for FIP200 P1, P2, P3, P4 and P6. The FIP200 P5 fragment did not amplify using the calculated annealing temperature; therefore, a temperature-gradient PCR was performed. A product of the expected size for FIP200 P5 was then observed as a distinct band, as shown in panel B.

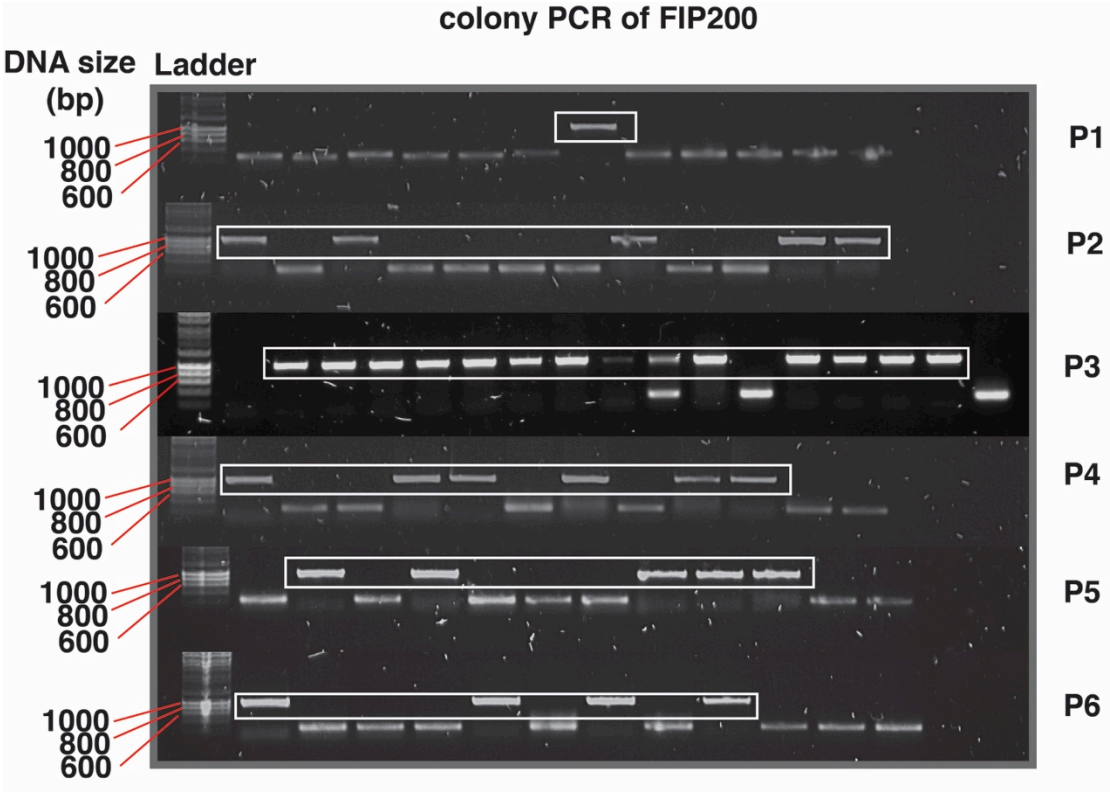


Figure 3.6: Agarose gels showing the results from the colony PCR used to identify plasmids containing the correct 800–1000 bp inserts corresponding to FIP200 (P1–P6). The plasmids from the colonies corresponding to the correct-sized bands were sequenced to verify ligation of correct DNA sequence.

Extensive expression trials were performed to identify the optimum conditions required for overexpression. After trying several *E. coli* strains, the six FIP200 constructs cloned into pCold™ TF expressed well as His_6-TF fusion proteins in BL21 (DE3) cells. Expression was verified on an SDS-PAGE gel, which showed discrete bands for the FIP200 constructs (P1–P6) at around 80 kDa, as expected for the His_6-TF-protein fragment fusions (Figure 3.4A). Large-scale expression was carried out using the optimised conditions, and SDS-PAGE analysis of His_6-TF-FIP200 fragments showed good expression levels (Figure 3.4C).

The His_6-TF-FIP200 (P1–P6) proteins were purified successfully by affinity chromatography using Ni-NTA resin following large-scale (1 L) expression. Removal of the His_6-TF region from the FIP200 fragments by thrombin cleavage resulted in the precipitation of the protein. Therefore, all the FIP200 fragments were used as His_6-TF

fusions for further studies. The protein purified using Ni-NTA resin was analysed on an SDS-PAGE gel (Figure 3.4D), which confirmed the purity of the protein samples obtained from all six FIP200 constructs.

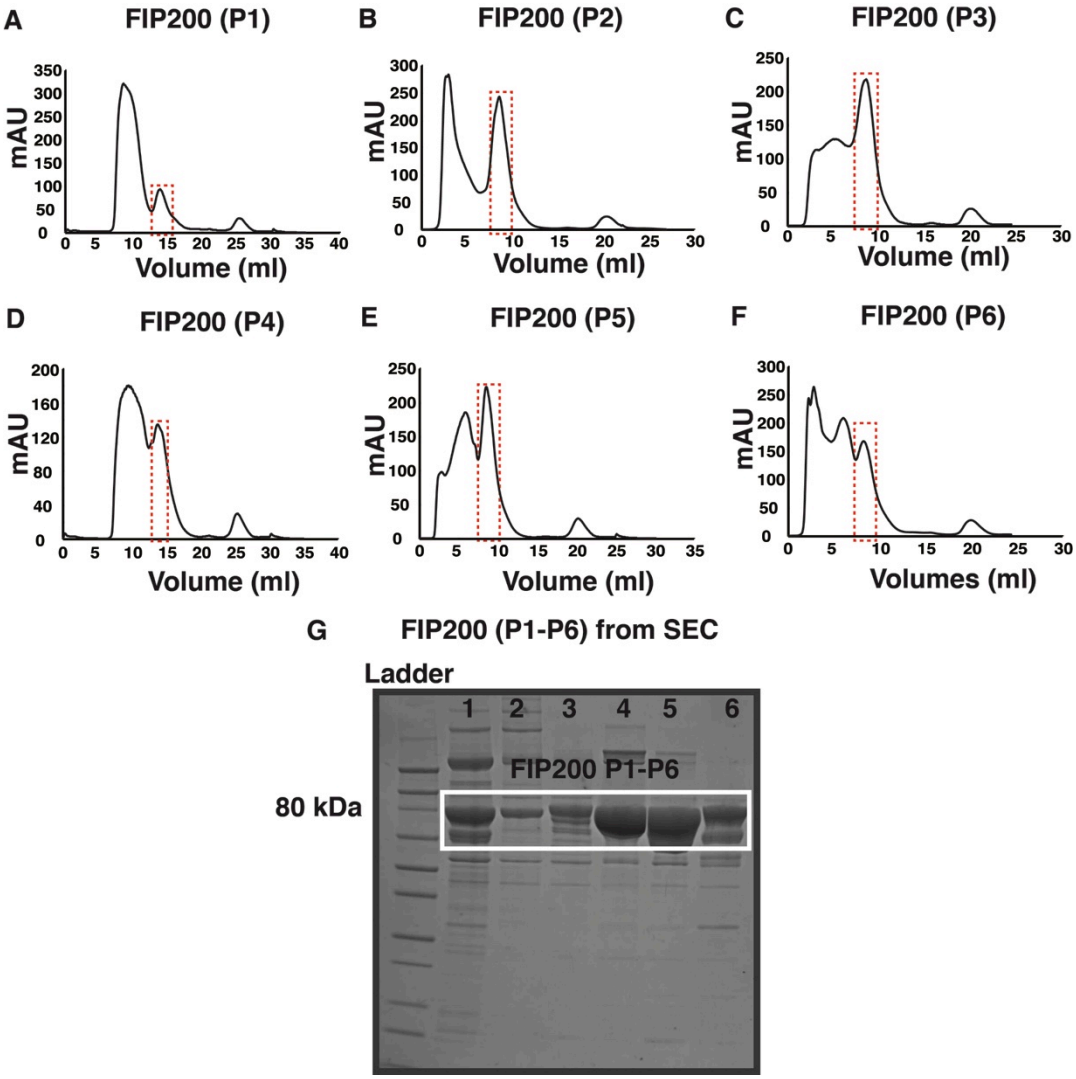


Figure 3.7: A–F. Chromatograms from SEC purification of the six FIP200 (P1–P6) fragments show more than one peak, suggesting that the proteins were not of sufficient purity for crystallisation. His_6-TF-FIP200 (P1–P6) constructs were purified, but were difficult to separate because of the unstable nature of the FIP200 (P1–P6) constructs. G. SDS-PAGE gel showing the fraction that had the FIP200 P1–P6 constructs. Results suggested that, despite observing bands of the correct size, the samples were not pure and were degraded, which was evident from the ladder-like bands in each lane. Lanes 1–6: FIP200 P1–P6.

To obtain pure protein samples for the crystallisation experiments, size exclusion chromatography was used as a final purification step after elution of all six His_6-TF-FIP200 fragments from the Ni-NTA resin. The fractions were collected in 1-mL volumes, and SDS-PAGE analysis (Figure 3.7G) showed peaks corresponding to

the FIP200 fragments. Several peaks were observed during purification by SEC (Figure 3.7A–F), but it was difficult to separate and purify the protein samples because of the unstable nature of the FIP200 fragments.

3.2.4 Pull-down assay demonstrated no binding between the ATG13 (380–517) construct and the FIP200 (P1–P6) constructs

Pull-down assays did not show the binding region of C-terminal ATG13 in FIP200. To identify the FIP200 (P1–P6) fragment that binds to the C-terminal of ATG13 (380–517), a GST pull-down assay was performed. The ATG13 (380–517)-GST fusion protein was used as "bait" by attaching it to glutathione Sepharose resin. The six His_6-TF-FIP200 (P1–P6) fragments were used as "prey" (refer to Section 2.5.6 and Figure 2.1). All six His_6-TF-FIP200 (P1–P6) fusion proteins were eluted from the Ni-NTA resin, and were then incubated with ATG13 (380–517)-GST attached to the glutathione Sepharose resin. After several rounds of washing, the incubated samples were loaded onto an SDS-PAGE gel to observe the binding. GST without any fusion protein was used as a control to ensure that any observed interaction was between the FIP200 fragments and ATG13, not with GST (Figure 3.8).

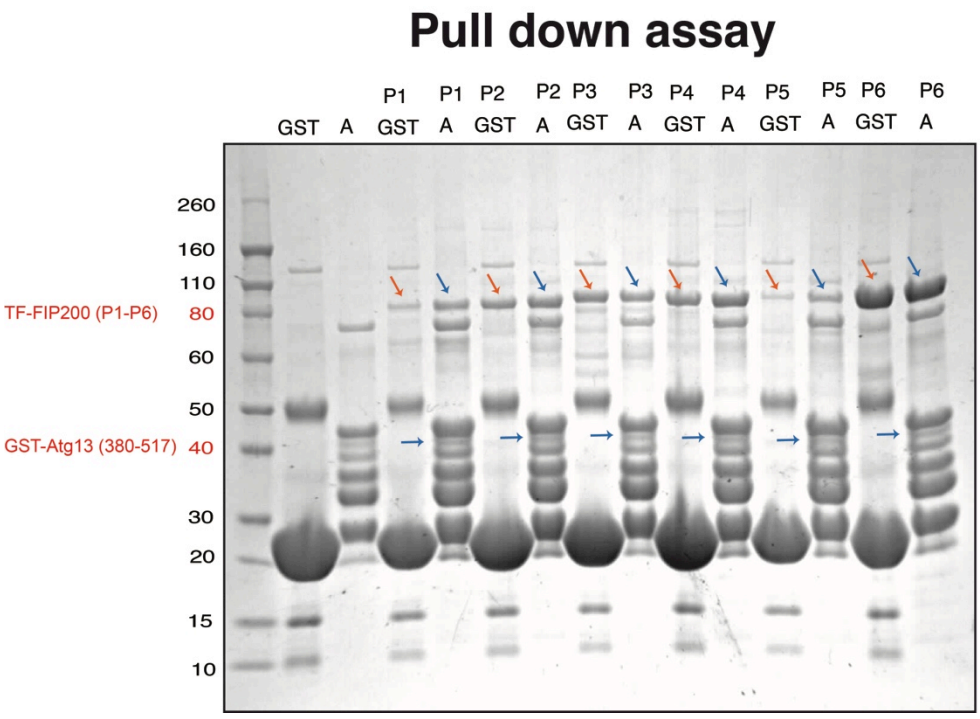


Figure 3.8: SDS-PAGE analysis of the pull-down assay demonstrated that there was no binding between ATG13 (380–517)-GST and the six His_6-TF-FIP200 (P1–P6) fragments. Molecular weight marker (Novex® Sharp Prestained Protein Standard) was loaded in lane 1. The GST control protein was run in lane 2, and ATG13 (380–517)-GST was run in lane 3. His_6-TF-FIP200 (P1–P6) fragments incubated with the GST control protein were run in lanes 4, 6, 8, 10, 12 and 14. His_6-TF-FIP200 (P1–P6) fragments incubated with ATG13 (380–517)-GST were run in lanes 5, 7, 9, 11, 13 and 15.

Interaction between Atg13 and FIP200 would be expected to show as bands on the SDS-PAGE gel at 80 kDa for FIP200 fragments and 40 kDa for ATG13 in the P1A–P6A lanes (Figure 3.8 shown in blue arrows). However, it was observed that all six His_6-TF-FIP200 polypeptides (P1–P6) interacted with the GST control (lanes P1 GST–P6 GST, shown as orange arrows in Figure 3.8). Therefore, it was difficult to distinguish whether the interaction of FIP200 polypeptides was with the ATG13 or with the GST region of the fusion protein. Another problem was the massive degradation of ATG13 (380–517)-GST attached to the GST resin, which occurred only upon binding of this protein to the GST resin, but not to sample eluted from the GST resin. Although interaction between FIP200 and Atg13 has been previously reported (Hosokawa *et al.*, 2009), it was not possible to identify the region of FIP200 that binds ATG13 using this pull-down assay because of the abovementioned reasons.

Extensive unsuccessful attempts were made to optimise the pull-down assay to avoid the false positive interaction of GST with FIP200 and overcome ATG13 (380–517)-GST degradation. However, because of the repeated failure of these attempts, the experiment was not pursued further.

3.2.5 Initial crystallisation trials on the FIP200 (P1–P6) constructs

Initial crystallisation experiments for the six His_6-TF-FIP200 (P1–P6) fragments did not result in any protein crystals. Protein samples from SEC purification were used for crystallisation experiments. Peak fractions corresponding to the expected size of the His6-TF-FIP200 (P1–P6) fragments were pooled and concentrated using 30,000 MWCO spin concentrators. Each fragment was concentrated to 10 mg/mL. Despite degradation of the protein samples and the samples not being sufficiently pure (Figure 3.7G) for crystallography experiments, initial crystallisation trials were still performed using Crystal Screen, Crystal Screen Cryo, Index and PEG/Ion (All from Hampton Research) and JCSG+, Midas, Morpheus, PACT, Clear Screen Strategy 1 and Clear Screen Strategy II (All from Molecular Dimensions). In all cases, the drops included 400 nL of protein and 400 nL of mother liquor. All crystallisation experiments were attempted at both room temperature and 4°C. No crystal hits were found in the initial trial, and further trials were not conducted due to the unstable nature of the proteins.

3.3 Summary

Although the region of ATG13 that binds to FIP200 is known, the ATG13-binding region in FIP200 remains elusive. To map the ATG13-binding region of FIP200, the full length FIP200 gene was truncated into six fragments, and was

successfully cloned into pCold™ TF vector and expressed in *E. coli* BL21 (DE3) cells. The ATG13 DNA sequence encoding amino acids 380–517 was successfully cloned into pGEX-4T-1 vector and was expressed in Rosetta (DE3) cells.

It was hypothesised that the previously reported C-terminal region of ATG13 would bind to one the six fragments of FIP200 (P1–P6), and the amino acids of the interacting fragment that are involved in binding could be further mapped. However, because of the false positive interaction between FIP200 polypeptides and the GST tag, the pull-down assay did not provide any useful information that could be used to map the interaction region in FIP200. It is possible that the FIP200 polypeptides were not properly folded, which caused them to interact with the GST protein. Using a different fusion partner for ATG13, such as maltose binding protein, and experimenting with different FIP200 fragments may be viable future approaches to troubleshoot the abovementioned problems.

In conclusion, despite the failure to map the region of FIP200 that interacts with ATG13, this chapter provided key information for the failure of the pull-down assay. Further optimisation based on the suggestions above may allow discovery of the binding region and mapping of the amino acids involved in the interaction.

Chapter 4. STRUCTURAL ANALYSIS OF THE LIR-DEPENDENT INTERACTION BETWEEN THE PLEKHM1 LIR AND HUMAN ATG8-LIKE PROTEINS.

4.1 Introduction

In the previous chapter, the interaction between FIP200 and ATG13, which are part of the ULK1 complex, was studied using pull-down assays. Autophagy-related protein 8 (Atg8) proteins are essential for the formation of the autophagosome (Nakatogawa *et al.*, 2007, Weidberg *et al.*, 2011, Xie *et al.*, 2008). Atg8 proteins form a multifunctional family that act as scaffolds for large complexes, and as adaptors that link the cargo to be degraded to the autophagosome (Nakatogawa *et al.*, 2007, Kirkin *et al.*, 2009, Novak *et al.*, 2010, Bjørkøy *et al.*, 2005, Shvets *et al.*, 2008). Although lower eukaryotes, such as yeast, express only one Atg8 protein (Lang *et al.*, 1998, Kirisako *et al.*, 1999), humans have six homologues that are subdivided into two families: the MAP1LC3 family (including LC3A, LC3B and LC3C), and the GABARAP family (including GABARAP, GABARAPL1 and GABARAPL2) (Shpilka *et al.*, 2011). In this chapter, the interactions of human ATG8 family proteins with pleckstrin homology domain-containing protein family member 1 (PLEKHM1) are studied.

PLEKHM1 is a multivalent adaptor protein that regulates the degradation of protein aggregates through either the selective or the non-selective autophagic

pathways. It facilitates fusion of the autophagosome with the lysosome, with the help of Atg8 family proteins (McEwan *et al.*, 2015). PLEKHM1 interacts with all members of the human ATG8 protein family, but with varying binding affinities (data not published, **Table 4.1**). In this chapter, X-ray crystallographic experiments were performed to rationalise the differences in binding affinities of members of the LC3 and GABARAP families of human ATG8 proteins with PLEKHM1, and to probe the mechanism of interaction.

PLEKHM1 regulates the selective degradation of defective ribosomal products and the removal of puromycin-induced aggresome-like induced structures in an LIR-dependent manner (Szeto *et al.*, 2006). *PLEKHM1* mutations cause osteoporosis (Van Wesenbeeck *et al.*, 2007) as a result of dysfunctional ruffled border formation, which is a specialised process that requires lysosome fusion and facilitates degradation of the bone matrix (Van Wesenbeeck *et al.*, 2007, Zhao *et al.*, 2001, DeSelm *et al.*, 2011). Understanding the interaction mechanism of PLEKHM1 could provide information about the cause of this disease.

The interaction of PLEKHM1 with mammalian Atg8 (mAtg8) is mediated through the LC3-interacting region (LIR) motif (Alemu *et al.*, 2012). A typical LIR motif sequence is W/Y/FxxL/I/V, where the first residue is an aromatic amino acid, followed by acidic residues (shown as "x"), with a hydrophobic residue at the end. Tryptophan and phenylalanine are the most common aromatic residues in LIR motifs, especially in the selective autophagy adaptors (Alemu *et al.*, 2012). The PLEKHM1 LIR has the sequence 634-EWVNV-638, with a tryptophan in the "W-site" and valine in the "L-site" (McEwan *et al.*, 2015). To date, the structures of various selective adaptors with mammalian Atg8 proteins (NDP52-LC3C, OPTN-LC3B, p62-LC3B, NBR1-GABARAPL1) have been solved (Ichimura, Kumanomidou*, et al.*, 2008,

Kirkin *et al.*, 2009, Novak *et al.*, 2010, Pankiv *et al.*, 2007, Suzuki *et al.*, 2014, Thurston *et al.*, 2009, von Muhlinen *et al.*, 2012a). Although, considerable effort has gone into understanding the LIR-dependent interaction between the adaptors and various human ATG8 modifiers, there are no crystal structures of a single LIR with all of the members of human ATG8 proteins, which could be used to explore how the LIR interactions differ between LC3 and GABARAP proteins.

The LC3 and GABARAP family proteins have similar crystal structures. Both have two domains: an N-terminal domain that consists of the α_1 and α_2 helices, and the C-terminal domain that has a ubiquitin-like form with a β-grasp fold (Figure 4.1A) (Sugawara *et al.*, 2004, Bavro *et al.*, 2002, Knight *et al.*, 2002). The same domain architecture was found in all five crystal structures of human ATG8 proteins in complex with the PLEKHM1 LIR (Figure 4.1A). Both the LC3 and GABARAP family proteins have two hydrophobic pockets: the "W-site" and the "L-site" (Figure 4.1B). In the interaction between the PLEKHM1 LIR and the LC3 and GABARAP family proteins, Trp635 docks into the "W-site" and Val638 docks into the "L-site" (Figure 4.1B).

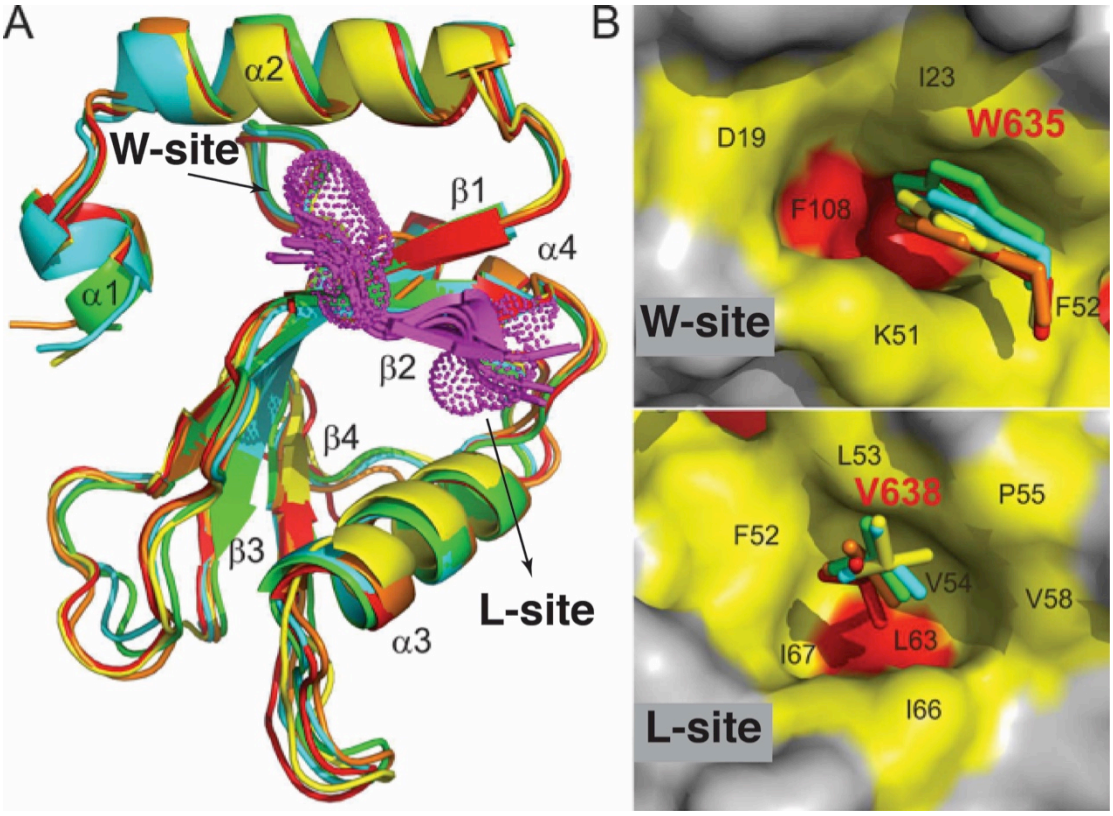


Figure 4.1: A. Structural overlay of PLEKHM1 LIR-LC3A (yellow), PLEKHM1 LIR-LC3B (orange), PLEKHM1 LIR-GABARAP (green) and PLEKHM1 LIR-GABARAPL1 (cyan) demonstrates that the domain architecture of the four human ATG8 protein structures is very similar. The PLEKHM1 LIR (shown in magenta) interacts with all of the human ATG8 structures in a similar orientation. The hydrophobic amino acids Trp635 and Val638, with the "W-site" and "L-site" hydrophobic surfaces, are shown as dot surface representations. B. The interaction of Trp635 from the PLEKHM1 LIR with the "W-site" hydrophobic surface (show as yellow surface representation) shows subtle differences in the orientation of Trp635 between the ATG8-family members, whereas there is no noticeable difference in the orientation of Val636 at the "L-site".

A recent report (McEwan *et al.*, 2015) demonstrated that the PLEKHM1 LIR binds all members of the human ATG8 protein family, but whether the PLEKHM1 LIR has a preference for a particular member of this protein family is unknown. To investigate this, the binding affinities of the PLEKHM1 LIR with all human ATG8 family proteins were determined using isothermal titration calorimetry (ITC) in Professor Ivan Dikic's laboratory. ITC data was kindly provided by Vladmir Rogov (**Table 4.1**).

Table 4.1: Thermodynamic parameters of the interaction between ATG8 family proteins and the PLEKHM1 LIR peptide.

	ΔH kcal mol^{-1}	**ΔS** cal mol^{-1} K^{-1}	**−T*ΔS** kcal mol^{-1}	**ΔG** kcal mol^{-1}	**K_a** $*10^6$ M^{-1}	**K_d** μM	**N**
LC3A	−7±0.2	+1.17	−0.35	−7.33	0.24±0.02	4.22	0.98±0.02
LC3B	−5.8±0.2	+4.27	−1.27	−7.09	0.16±0.01	6.33	1.06±0.01
LC3C	−8.3±0.2	−2.83	+0.84	−7.48	0.29±0.02	3.45	0.99±0.02
GABARAP	−10.6±0.06	−6.92	+2.06	−8.54	1.8±0.1	0.55	1.00±0.01
GABARAP L1	−7.8±0.1	+1.94	−0.58	−8.35	1.3±0.1	0.77	1.00±0.01
GABARAP L2	−6.14±0.03	+7.00	−2.09	−8.23	1.07±0.03	0.93	1.05±0.01

These results show that GABARAP family proteins have higher binding affinity towards the PLEKHM1 LIR than LC3 family proteins. GABARAP (K_D = 0.55 μM) has the highest binding affinity towards the PLEKHM1 LIR, and LC3B (K_D = 6.3 μM) has the lowest binding affinity, a 12-fold difference. The difference in binding affinity between LC3C (K_D = 3.45 μM), which has the highest binding affinity among LC3 family proteins, and GABARAPL2 (K_D = 0.93 μM), which has the lowest binding affinity among GABARAP family proteins, is four-fold, indicating that all GABARAP family proteins have higher binding affinity towards the PLEKHM1 LIR than the LC3 family proteins.

A more detailed analysis of the interaction between the PLEKHM1 LIR and ATG8 family proteins may explain the differences in binding affinity. However, there is no structural information on a single LIR motif interacting with all members of both the LC3 and GABARAP family proteins. In this chapter, the crystal structures of the PLEKHM1 LIR peptide in complex with LC3A, LC3C, GABARAP and GABARPL1 were determined using X-ray crystallography. The structures were then compared to understand the structural basis of the differences in binding affinity observed using ITC. The results suggest that small differences in the interaction of the PLEKHM1 LIR with GABARAP proteins and LC3 proteins contribute to the varying binding affinity.

4.2 Results

4.2.1 Purification and crystallisation of the PLEKHM1 LIR-human ATG8 complexes.

To shed light on the differences in binding affinities between human ATG8 proteins and the PLEKHM1 LIR, crystal structures of human ATG8 proteins in complex with the PLEKHM1 LIR peptide were determined. Pure protein samples were required for crystallisation experiments. To achieve this, genes for each of the ATG8 proteins were cloned into the pET30ΔSE vector as chimeric constructs to express His_6-tagged PLEKHM LIR-ATG8 family fusion proteins. Proteins were expressed in *E. coli* BL21(DE3) cells. Following IMAC, SEC resulted in purification of the target protein samples, as shown by the sharp peaks in the chromatograms and on SDS-PAGE gels (Figure 4.2A–D).

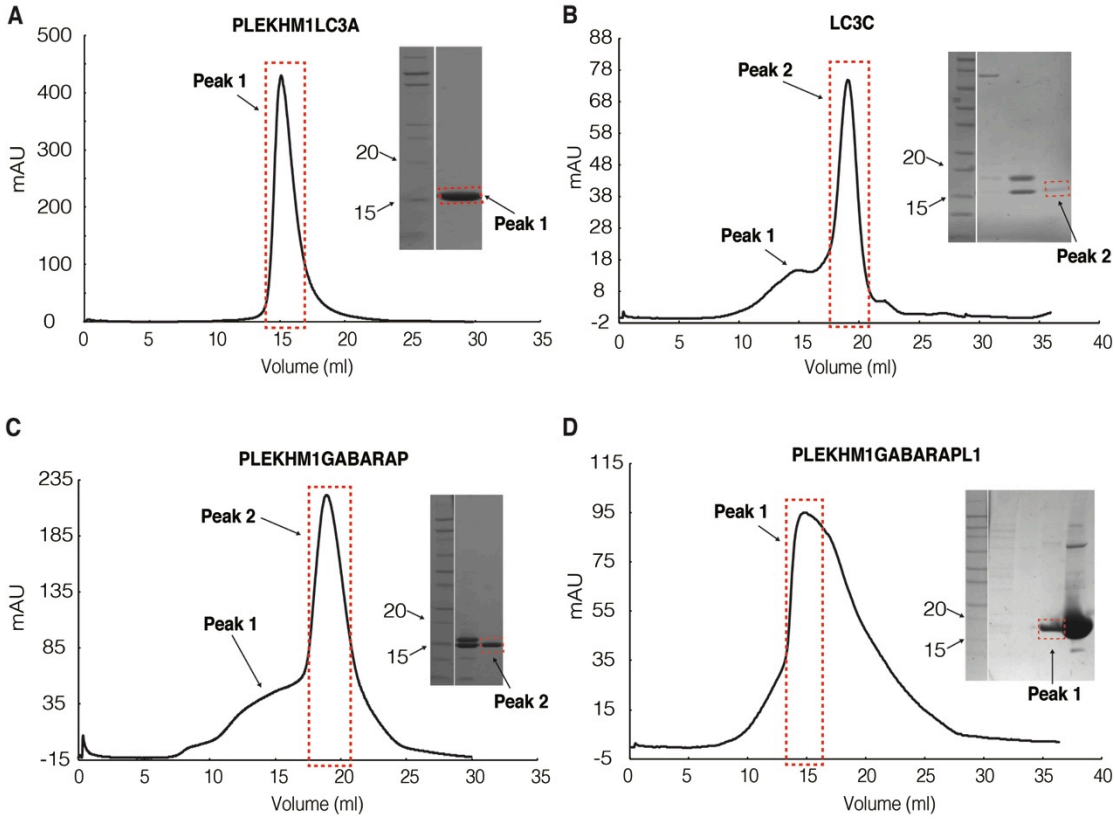


Figure 4.2: Purification of His_6-PLEKHM1-ATG8 family chimeric proteins by SEC. Peaks containing protein that was used for crystallisation are marked with red boxes. Inset: SDS-PAGE gels show purity of the samples obtained following SEC, and bands marked with red boxes show that the target protein separated at the correct size.

Full-length LC3C was unstable and was difficult to obtain in concentrations high enough for crystallisation. Thus, a truncated version of LC3C (8–128) was expressed with a His_6 tag. Vladmir Rogov encountered a similar problem, and expressed a truncated LC3C region (11–125) that was used in ITC experiments (Personal communication). LC3C (8–128) was purified using the same procedure as that used for the chimeric protein samples, except that the entire purification procedure was performed at 4°C, including the overnight thrombin cleavage for removal of the His_6 tag. Ultimately, pure target protein samples were successfully obtained and crystallised (Section 2.5.7.1).

4.2.2 Structural characterisation of the PLEKHM1 LIR in complex with members of the human ATG8 protein family demonstrates subtle differences in the crystal structures.

Given the very similar structure and high sequence similarity between members of the human ATG8 protein family, it was unclear why the PLEKHM1 LIR peptide interacts with the different proteins with different binding affinities. To understand this, the interactions between these proteins and the PLEKHM1 LIR were examined in atomic detail. The crystal structures of LC3A, LC3C, GABARAP and GABARAPL1 in complex with the PLEKHM1 LIR were determined, as described in Section 2.5.7. The chimeric protein crystals diffracted at 2.5 Å for PLEKHM1-LC3A, 2.0 Å for PLEKHM1-GABARAP and 2.9 Å for PLEKHM1-GABARAPL1. The LC3C (8–128) protein was co-crystallised with the PLEKHM1 LIR (629–642) peptide, and diffracted to 2.19-Å resolution. The crystal structure of the PLEKHM1-LC3B complex (PDB code: 3X0W) was recently solved (McEwan *et al.*, 2015), and was included for structural comparison. The data collection statistics and refinement statistics are summarised in Table 4.2.

Table 4.2: Summary of crystal structure data collection and refinement statistics

	PLEKHM1-GABARAP	**PLEKHM1-GABARAPL1**	**PLEKHM1-LC3A**
Data-collection statistics			
Space group	$P3_1$	$P3_221$	$P2_1$
Cell parameters a, b, c (Å); α, β, γ (°)	73.8, 73.8, 74.5; 90, 90, 120	84.8, 84.8, 105.4; 90, 90, 120	62.7, 63.1, 84.9; 90, 100.9, 90
Wavelength (Å)	0.9537	0.9537	0.9537
Resolution (Å)	37.23-2.00 (2.05-2.00)	42.8-2.90 (3.03-2.90)	45.57-2.60 (2.50-2.50)
R_{Merge}	0.196 (3.387)	0.148 (2.352)	0.113 (0.614)
$R_{p.i.m.}$	0.060 (1.066)	0.048 (0.735)	0.068 (0.366)
$R_{meas.}$	0.205 (3.652)	0.156 (2.466)	0.132(0.716)
I/σI	10.3 (1.5)	10.6 (1.4)	7.1 (2.1)
Completeness	99.2 (98.6)	99.9 (99.5)	100 (100)
CC(1/2)	0.998 (0.651)	0.991 (0.581)	0.928 (0.754)
Redundancy	11.6 (11.6)	10.9 (11.0)	3.8 (3.8)
Total reflections (unique)	353,770 (26,146)	110,277 (17,825)	22,780 (2,562)
Refinement statistics			
R_{Work}/R_{Free} (%)	18.4/22.3	21.4/26.0	21.8/26.1
Number of atoms			
Protein	2,990	1,949	4,086
Water	177	0	42
Temperature factors			
Main chain	34.4	100.8	54.7
Side chain	37.1	103.4	64.2
LIR peptide	30.6	110.4	50.7
Solvent	37.5	-	37.6
Ramachandran plot			
Allowed regions	100.0	99.6	99.4
Outliers	0.0	0.41	0.61
R.m.s.d from ideality			
Bond lengths (Å)	0.008	0.003	0.006
Planarity (°)	0.005	0.004	0.005
Angles (°)	1.040	0.778	1.077

Table 4.3: Summary of crystal structure data collection and refinement statistics for LC3C-PLEKHM1

	LC3C PLEKHM1 cocrystal	**LC3C PLEKHM1 cocrystal**
Data-collection statistics		
Space group	*P*1	$P2_1$
Cell parameters	54.1, 73.2, 89.1;	54.01, 89.15,
a, b, c (Å); α, β, γ (º)	89.96, 89.99,	73.53; 90,
	77.3	102.75, 90
Wavelength (Å)	0.9537	0.9537
Resolution (Å)	38.61-2.19	47.8-2.20
	(2.24-2.19)	(2.27-2.20)
R_{Merge}	0.074 (0.875)	0.053 (0.353)
$R_{p.i.m.}$	0.048 (0.573)	0.021 (0.138)
$R_{meas.}$	0.089 (1.063)	0.056 (0.379)
I/σI	13.4 (1.9)	22.3 (5.2)
Completeness	98.6 (97.8)	99.8 (99.2)
CC(1/2)	0.999 (0.824)	0.999 (0.974)
Redundancy	6.6 (6.6)	7.5 (7.5)
Total reflections (unique)	446,563	260,262
	(68,176)	(34,469)
Refinement statistics		
R_{Work}/R_{Free} (%)	20.0/24.6	23.0/30.78
Number of atoms		
Protein	8,300	4,109
Water	170	143
Temperature factors		
Main chain	49.8	46.34
Side chain	58.4	60.27
LIR peptide	62.4	64.75
Solvent	40.2	44.89
Ramachandran plot		
Allowed regions	99.1	98.92
Outliers	0.9	1.08
R.m.s.d from ideality		
Bond lengths (Å)	0.004	0.015
Planarity (°)	0.004	0.009
Angles (°)	0.732	1.334

The crystal structure of PLEKHM1-GABARAP was solved in the trigonal crystal system $P3_1$, and had three molecules in an asymmetric unit (Figure 4.3A). Each molecule interacted with the PLEKHM1 LIR region of a symmetry-related molecule (i.e. the LIR region of molecule A interacts with the symmetry-related molecule A). The PLEKHM1-GABARAPL1 structure was solved in a trigonal crystal

system, $P3_221$. Instead of three molecules, the PLEKHM1-GABARAPL1 crystal structure had two molecules in the asymmetric unit (Figure 4.3B). Unlike PLEKHM1-GABARAP, the LIR region in PLEKHM1-GABARAPLI of molecule A interacted with molecule B within the asymmetric unit, and the LIR region of molecule B interacted with symmetry-related molecule A (Figure 4.3B). The typical LIR hydrophobic residues Trp635, Val636 and Val638 of the PLEKHM1 LIR, which interact with the GABARAP and GABARAPL1 proteins, were clearly visible in the electron density map (Figures 4.3C and 4.3D).

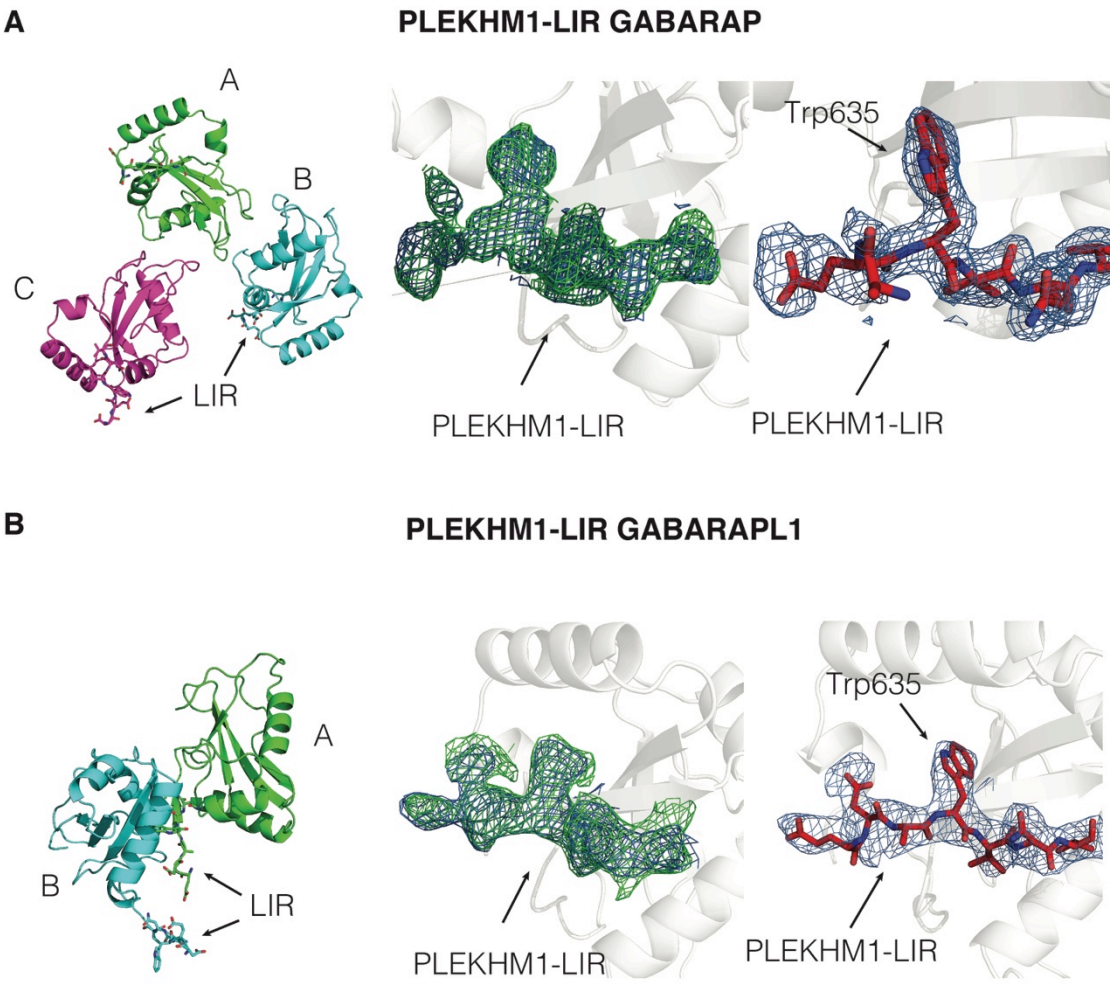


Figure 4.3: A. Left panel shows three monomers of PLEKHM1 LIR-GABARAP (labelled as A, B and C, and coloured by chains) found in an asymmetric unit, represented as a cartoon structure. Middle panel shows the OMIT map of the PLEKHM1 LIR. Right panel shows the 2Fo-2Fc map (blue) at 1 σ, the Fo-Fc map (green) at 3 σ and the Fo-Fc map (red) at −3 σ of the PLEKHM1 LIR, as was manually modelled. Clear density around the hydrophobic residues Trp635 and Val636, which are involved with the "W-site" and "L-site" hydrophobic surfaces, is also shown. B. Left panel shows two monomers of PLEKHM1 LIR-GABARAPL1 (labelled as A and B, and coloured by chains) found in an asymmetric unit, represented as a cartoon structure. Middle panel shows the OMIT map of the PLEKHM1 LIR. Right panel shows the 2Fo-2Fc map (blue) at 1 σ, the Fo-Fc map (green) at 3 σ and the Fo-Fc map (red) at −3 σ of the PLEKHM1 LIR, as was manually modelled. Clear density is also shown around residues Trp635 and Val636.

The crystal structure of PLEKHM1-LC3A was solved in the monoclinic crystal system $P2_1$, with four molecules in an asymmetric unit (Figure 4.5A). Unlike the LIR interactions observed in the PLEKHM1-GABARP and PLEKHM1-GABARAPL1 crystal systems, in the PLEKHM1-LC3A system, the LIR region of molecules A and B interacted with the symmetry-related molecules A and B, respectively, whereas the LIR region of molecule C interacted with symmetry-related

molecule D, and the LIR region of molecule D interacted with symmetry-related molecule C.

PLEKHM1-LC3C was integrated in space group $P2_1$, but refinement was unsuccessful because of the presence of translational pseudo-symmetry. The processed data set from XDS was scaled using *AIMLESS* (Evans, 2006) from the CCP4 software package. Analysis of systematically absent reflections determined 2_1 screw axes along *b*, confirming the $P2_1$ space group (Figure 4.4A). Therefore, the data was processed in space group $P2_1$, and a diagnostic test for translational pseudo-symmetry was performed using *phenix.xtriage* from the PHENIX suite (Adams *et al.*, 2010). Data were analysed from 10.00 Å to 2.84 Å resolution, and the intensity statistics suggested translational pseudo-symmetry with a value for $(I^2)/(I)^2$ of 2.667, which is not consistent with the theoretical value of 2. Additionally, the intensity probability distribution produced an abnormal curve, suggesting the presence of translational pseudo-symmetry (Figure 4.4B). However, the L-test did not show any twinning in the dataset (Figure 4.3C). Therefore, the data was solved and refined in the *P*1 space group (Table 4.3), and this model was used for molecular replacement. After iterative refinement and model building, the R_{free} did not go below 30%, and the difference between R_{free} and R_{work} remained high at 7% (Table 4.3). Therefore, the complex solved in the triclinic crystal system *P*1 was used for further analysis.

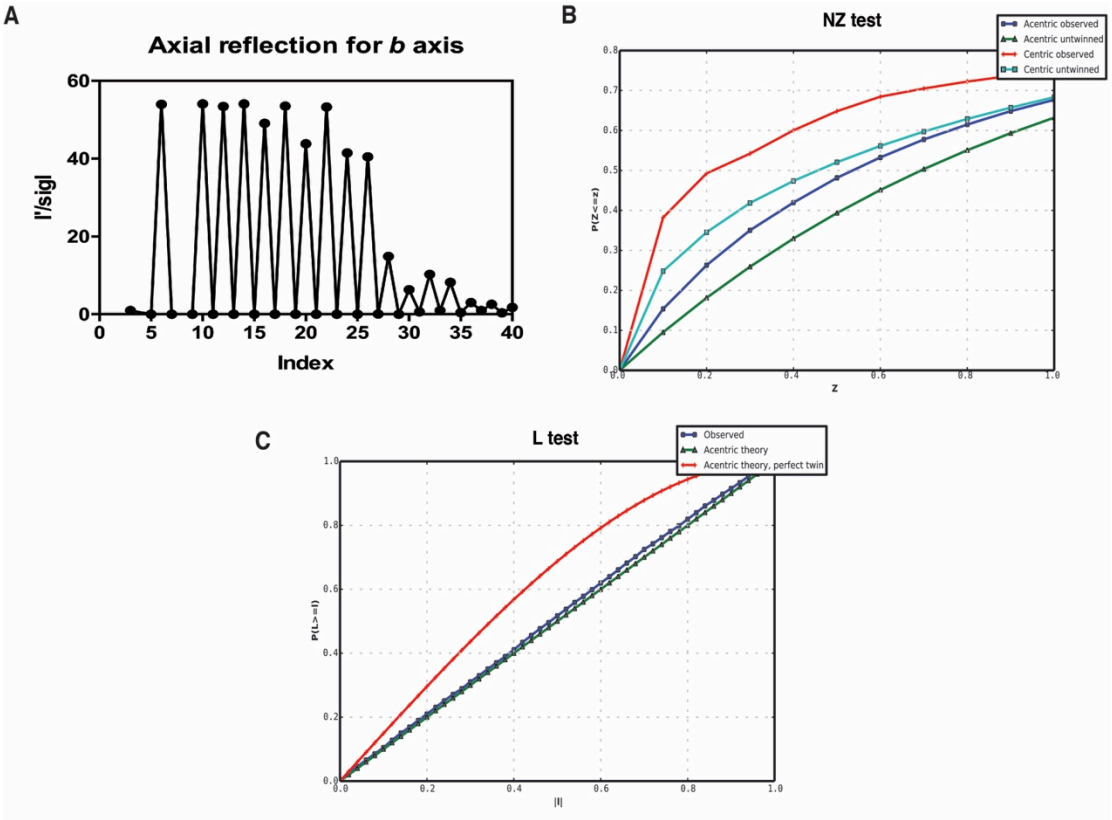


Figure 4.4: A. Axial reflection along the b axis with no violations from 40 measurements, suggesting a clear 2_1 screw axes along the b axis. B. Traditional cumulative intensity distribution normalised in resolution shells. The acentric observed and untwinned are shown in blue and green curves. The centric observed and untwinned are shown in red and blue curves. The abnormal distribution of the curves shows a clear translational pseudo-symmetry. C. The cumulative probability distribution of the local intensity statistics is shown, where the acentric theory (green) and the observed (blue) are compared with perfectly twinned data (red). The dataset shows no twinning.

The asymmetric unit (*P*1 space group) had eight monomers (Figure 4.5B), and each monomer had one molecule of LC3C interacting with one molecule of the PLEKHM1 LIR peptide. The electron density for the PLEKHM1 LIR peptide was clearly visible, and the amino acids were assigned according to the PLEKHM1 LIR peptide sequence (Fig 4.5C and 4.5D). Overall the data collection statistics and the refinement statistics of all four crystal structures suggest that the models are reliable, and can therefore be used for further structural analysis.

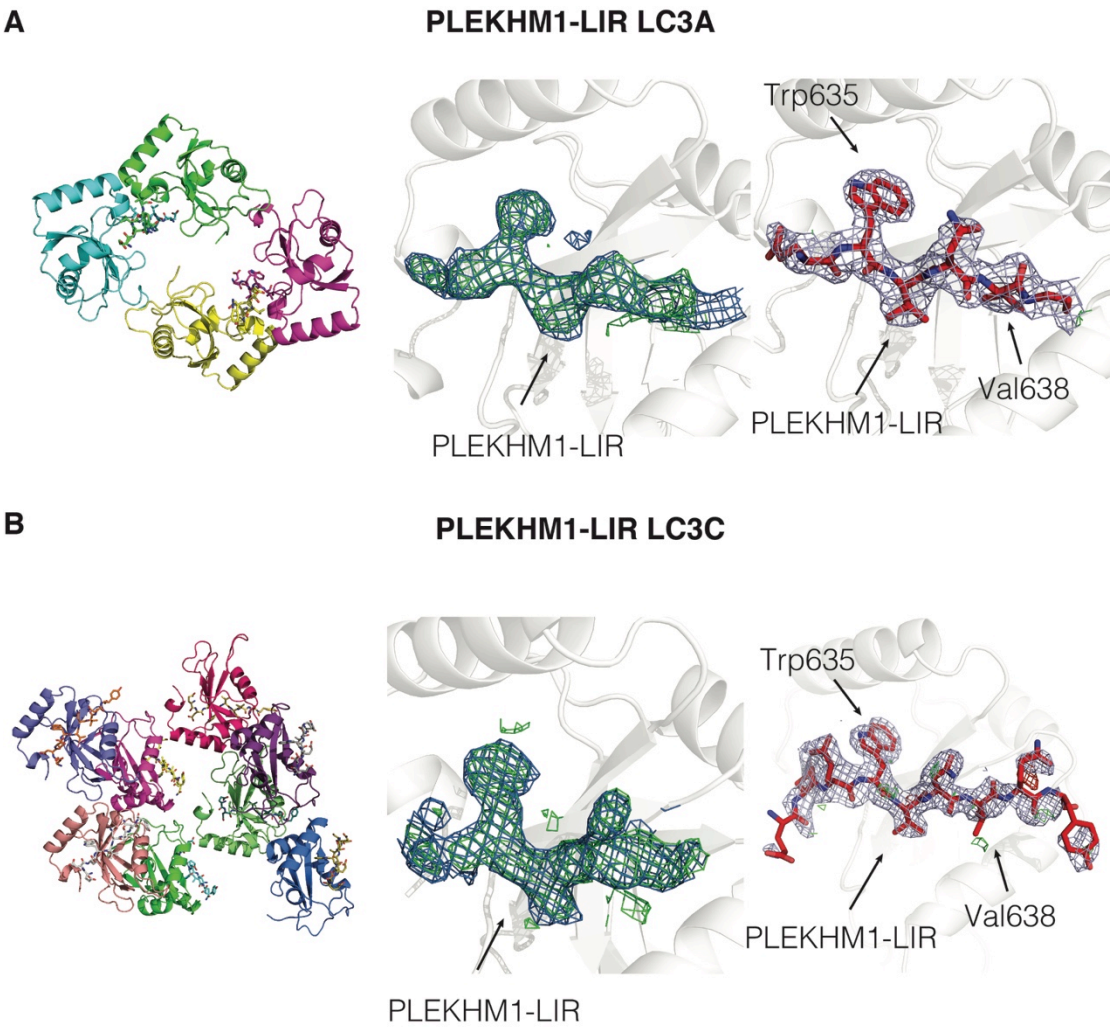


Figure 4.5: A. Left panel shows four monomers of PLEKHM1 LIR-LC3A (coloured by chains) found in an asymmetric unit, represented as a cartoon structure. Middle panel shows the OMIT map of the PLEKHM1 LIR. Right panel shows the 2Fo-2Fc (blue) at 1 σ, Fo-Fc (green) at 3 σ and Fo-Fc (red) at −3 σ maps of the PLEKHM1 LIR, which was manually modelled. It also shows clear density around the hydrophobic residues Trp635 and Val636, which are involved with the "W-site" and "L-site" hydrophobic surfaces. B. Left panel shows two monomers of PLEKHM1 LC3C (coloured by chains) found in an asymmetric unit, represented as a cartoon structure. Middle panel shows the OMIT map of the PLEKHM1 LIR. Right panel shows the 2Fo-2Fc (blue) at 1 σ, Fo-Fc (green) at 3 σ and Fo-Fc (red) at −3 σ maps of the PLEKHM1 LIR, which was manually modelled. It also shows clear density around the hydrophobic residues Trp635 and Val636.

The monomers in an asymmetric unit of each crystal structure were overlaid/aligned (Figures 4.6A, B, C & D) to explore the binding heterogeneity. The overlays show that the LIR region (EWVNV) of the PLEKHM1 LIR peptide was in largely the same orientation across all the monomers in all five crystal structures (including PLEKHM1-LC3B, PDB code: 3X0W, also included in the analysis). The

monomers in an asymmetric unit that contains the PLEKHM1 LIR fused at the N-terminal were overlaid, and the root-mean-square deviation (RMSD) for the α-carbon was calculated. The monomers overlaid with RMSDs of 0.174 Å for PLEKHM1-LC3A, 0.217 Å for PLEKHM1-LC3B, 0.167 Å for PLEKHM1-GABARAP and 0.326 Å for PLEKHM1-GABARAPL1 (Figure 4.6A–E). For the PLEKHM1-LC3C complex, the LC3C molecule interacting with the PLEKHM1 peptide was taken as one monomer and was overlaid with the other seven monomers (LC3C in complex with PLEKHM1 peptide), and displayed a RMSD of 0.426 Å. The overlays of the PLEKHM1 LIR showed fluctuation in the amino acid residues present on either side of the PLEKHM1 core LIR region (WVNV), but these residues aligned well across all of the monomers (Figure 4.6F). This suggested strong binding between hydrophobic residues of the PLEKHM1 LIR and the "W-site" and "L-site" hydrophobic surfaces of human ATG8 proteins. The RMSD values suggest that the monomers overlay very well. Therefore, subsequent analyses, performed using monomer A, should be consistent across all the monomers.

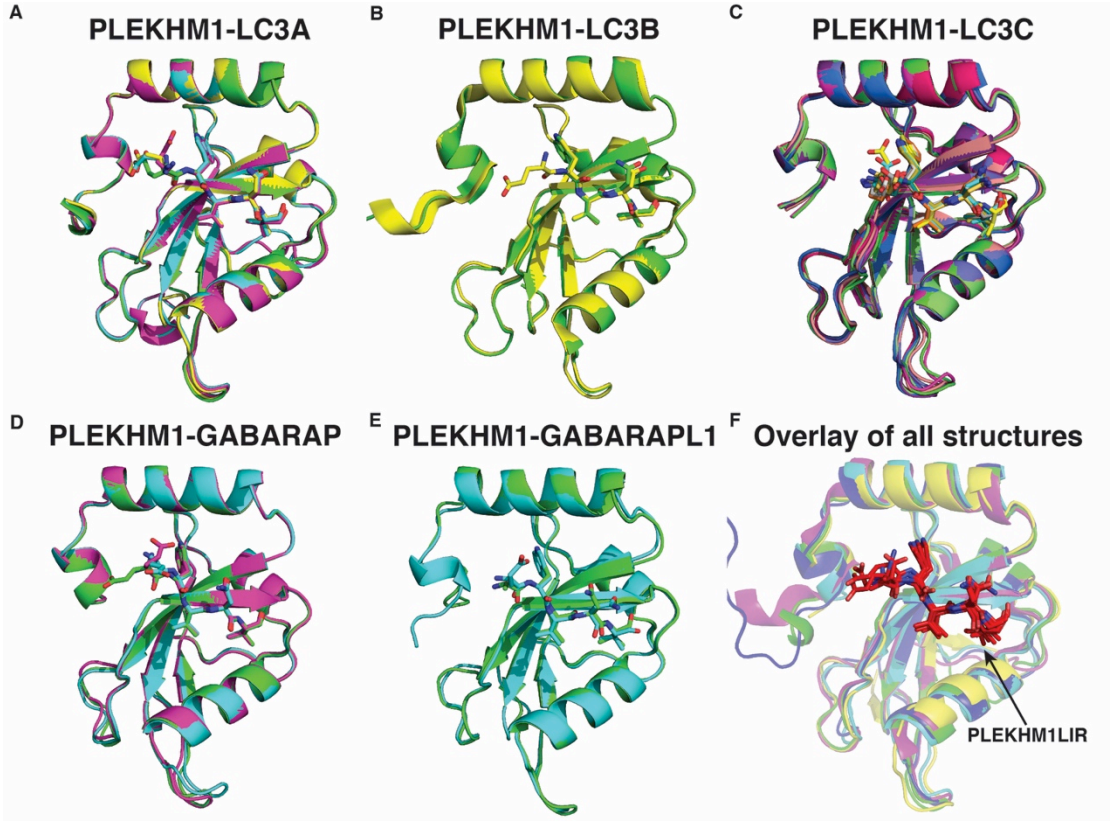


Figure 4.6: A. Overlay of the four PLRKHM1-LC3A monomers in the asymmetric unit. The monomers overlay well, and the PLEKHM1 LIR, represented as sticks, is in the same orientation in all four monomers, with the key amino acid Trp635 aligned well. B–E. Overlay of monomers in an asymmetric unit for PLEKHM1-LC3B, PLEKHM1-LC3C, PLEKHM1-GABARAP and PLEKHM1-GABARAPL1 structures, respectively. The monomers and the PLEKHM1 LIR overlay well. F. Overlay of all five human ATG8 protein structures. One of the monomers in each asymmetric unit from A–E was used for analysis The PLEKHM1 LIR is represented as red sticks.

Crystal structures of LC3A, LC3B, LC3C, GABARAP and GABARAPL1 displayed similar architecture to other ATG8 protein structures, with two domains (N-terminal and C-terminal) (Fig 4.1A). A typical LIR interaction was observed between the PLEKHM1 LIR and the human ATG8 proteins, where Trp635 of the PLEKHM1 LIR interacts with the hydrophobic surface at the "W-site" of the human ATG8 family protein, and Val638 of the PLEKHM1 LIR interacts with the hydrophobic surface at the "L-site" (Fig 4.1B). As observed in the other LIR interactions, the main chains of Val636 and Val638 of the PLEKHM1 LIR form hydrogen bonds with Lys51 and Leu53 in LC3A and LC3B, Lys57 and Leu59 in LC3C and Lys48 and

Leu50 in GABARAP and GABARAPL1 (Figure 4.7), in a β-strand like conformation.

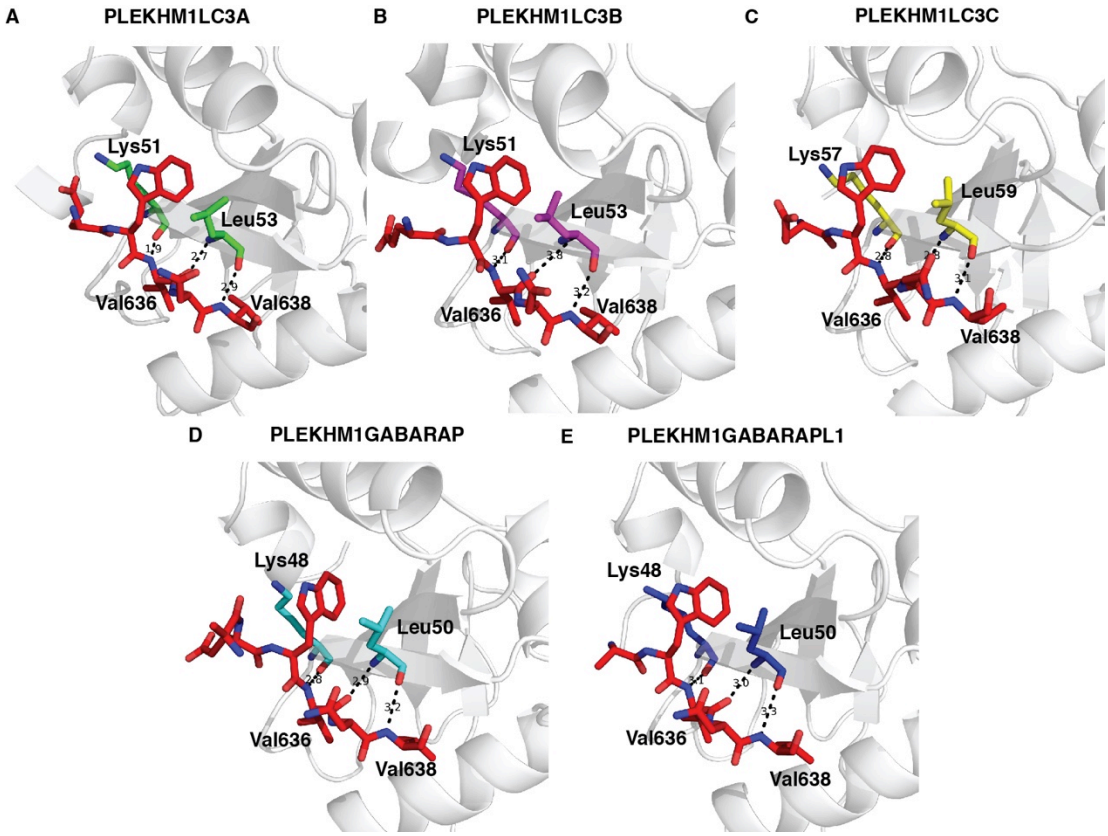


Figure 4.7: Comparison of β_2-strand hydrogen bond interaction with the PLEKHM1 LIR. The figure depicts the interaction of Val636 and Val638 (PLEKHM1 LIR shown in red) with the β_2 strand of ATG8 family proteins. A. Lys51 and Leu53 of LC3A, shown in green, make hydrogen bonds with Val636 and Val638 of the PLEKHM1 LIR, represented as red sticks. B–E. LC3B residues are shown in magenta, GABARAP residues in cyan, LC3C residues in yellow and GABARAPL1 residues in blue.

4.2.3 The interaction of human ATG8 proteins with Trp635 and Asn637 of the PLEKHM1 LIR suggests a tighter interaction of GABARAP and GABARAPL1 with the PLEKHM1 LIR.

Two pairs of interactions observed in all five structures are summarised in **Table 4.4**. Structural analysis suggests that tighter interaction between Arg28 and Asn637, coupled with the hydrogen bond between Glu17 and Trp635, may contribute to the higher binding affinity of GABARAP and GABARAPL1 towards the

PLEKHM1 LIR, when compared with the members of LC3 protein. The distance between Asp19 and Trp635, and Lys30 and Asn637, in the PLEKHM1 LIR-LC3B complex suggests LC3B has the weakest interaction with the PLEKHM1 LIR among all the human ATG8 homologues, which is consistent with the ITC data.

Table 4.4: Summary of the interaction of Trp635 and Asn637 with human ATG8 family members.

	Arg28-Asn637 (GABARAPs) Lys30-Asn637 (LC3A and LC3B) Lys36-Asn637 (LC3C)	**Glu17-Trp635 (GABARAPs) Glu19-Asn637 (LC3A) Asp19-Asn637 (LC3B) Glu25-Asn637 (LC3C)**	**Binding Affinity K_D (μM)**
GABARAP	2.8 Å	3.6 Å	0.55
GABARAPL1	3.1 Å	3.4 Å	0.77
LC3C	3.1 Å	No interaction	3.45
LC3A	No interaction	3.5 Å	4.22
LC3B	No interaction	No interaction	6.33

All five PLEKHM1-human ATG8 complexes had very similar structures, and the orientation of the PLEKHM1 LIR was similar in all cases. Given the differences in binding affinities for the PLEKHM1 LIR between the human ATG8 proteins, we hypothesized that subtle differences in the interaction between these molecules could contribute to the differences in binding affinity. To evaluate this hypothesis, the amino acids involved in the interactions of the PLEKHM1 LIR with ATG8 homologues were carefully analysed. The analysis showed that a lysine residue was located in the β_2-sheet region of LC3A (Lys30), LC3B (Lys30) and LC3C (Lys36), whereas Arg28 was observed at the same position in GABARAP and GABARAPL1 (Figure 4.7).

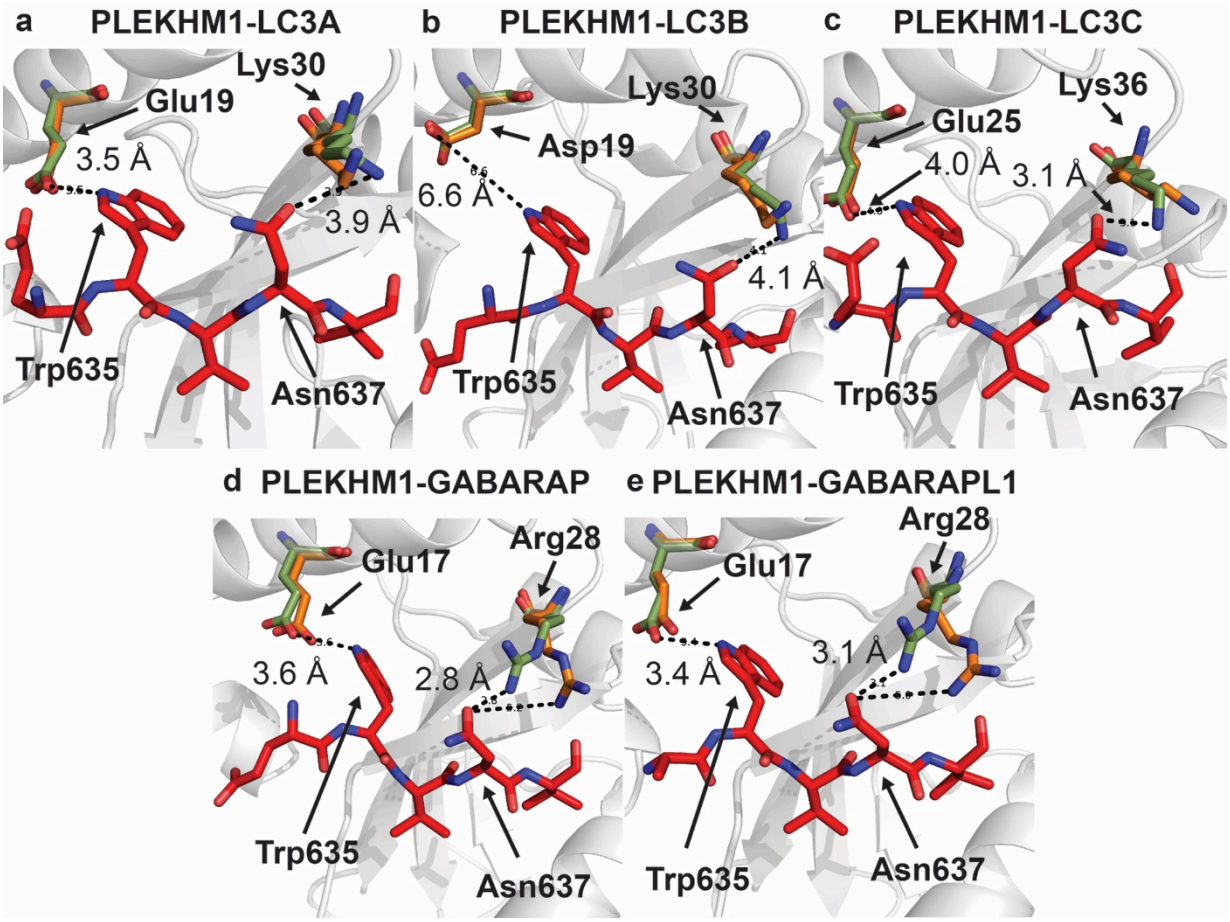


Figure 4.8: Interaction of Trp635 and Asn637 with ATG8 family members. The residues in green represent the PLEKHM1 LIR bound state, and the residues in orange represent the unbound state. The PLEKHM1 LIR peptide is shown in red. A. Interaction of the PLEKHM1 LIR with LC3A. Trp635 of PLEKHM1 interacts with Glu19 of LC3A (3.5 Å distance). Asn637 of PLEKHM1 interacts with Lys30 (3.9 Å distance). B. LC3B interacts in a similar way to LC3A, but the bond distances are higher (6.6 Å and 4.1 Å, respectively), suggesting a weaker interaction. C. LC3C also interacts in a similar way to LC3A, but forms a strong interaction between Asn637 and Lys36 (3.1 Å distance). D. Trp635 forms a strong interaction with Glu17 in GABARAP (3.5 Å distance). Instead of Lys interacting with Asn637, GABARAP proteins interact through Arg28, which flips around 4 Å between bound and unbound states, and forms a tight interaction with Asn637 (2.8 Å distance). E. GABARAPL1 interacts in a similar way to GABARAP, with Asn637 interacting with Arg28 (3.1 Å distance).

Asn637 in the PLEKHM1 LIR forms a hydrogen bond with Arg28 of GABARAP and GABARAPL1 (Figure 4.8D and E). A similar interaction at the same position was found in LC3C, between Lys36 and Asn637 of the PLEKHM1 LIR (Figure 4.7C), but the position of Lys30 in LC3A and LC3B made it difficult to interact with Asn637 (Figure 4.8A and B). In GABARAP and GABARAPL1, the comparison between the liganded and unliganded states shows that the side chain of

Arg28 undergoes a shift of 4 Å, and nitrogen ($N^{\eta 2}$) of Arg28 forms a hydrogen bond with the oxygen ($O^{\delta 1}$) of Asn637 at a distance of 2.8 Å for GABARAP and 3.1 Å for GABARAPL1 (Figure 4.8D and E). In LC3C, comparison between the liganded and unliganded states shows that, unlike Arg28 in GABARAP and GABARAPL1, the side chain of Lys36 does not change with respect to its ideal position. Instead, the nitrogen (N^{ζ}) forms a hydrogen bond with $O^{\delta 1}$ of Asn637 in the PLEKHM1 LIR at a distance of 3.1 Å (Figure 4.8 C). In LC3A and LC3B, comparison between the liganded and unliganded states shows that the side chain of Lys30 does not rearrange. Greater distances between $O^{\delta 1}$ of PLEKHM1 LIR Asn637 and N^{ζ} of Lys30 in LC3A (3.9 Å) and LC3B (4.1 Å) suggest that hydrogen bond formation between these two residues was more difficult (Fig 4.8A and B). Therefore, this analysis suggests that the hydrogen bond formed between GABARAP and Asn637 of PLEKHM1 is the tightest among all five PLEKHM1 LIR-human ATG8 structures, followed by GABARAPL1 and LC3C.

The second difference between the five PLEKHM1 LIR-human ATG8 complex structures was the weak hydrogen bond formed between $O^{\varepsilon 2}$ of Glu17 in GABARAP and GABARAPL1 and $N^{\varepsilon 1}$ of Trp635 in the PLEKHM1 LIR (Figure 4.8D and E). A similar interaction was found between $O^{\varepsilon 2}$ of Glu19 in LC3A and $N^{\varepsilon 1}$ of Trp635 of the PLEKHM1 LIR (Figure 4.8A). In the PLEKHM1 LIR-LC3C complex, the distance between corresponding amino acid Glu25 and Trp635 was 4.5 Å, which made hydrogen bond formation difficult (Figure 4.8C). Unlike other human ATG8 family proteins, the corresponding position in LC3B is an aspartic acid (Asp19), rather than a glutamic acid. Aspartic acid has a shorter side chain than glutamic acid, and therefore the distance between Asp19 of LC3B and Trp635 of the PLEKHM1 LIR (6.6 Å) is less suitable for hydrogen bond formation (Figure 4.8B).

Overall, these interactions provide a logical reason for the varying binding affinities between the PLEKHM1 LIR and human ATG8 family proteins.

4.2.4 Analysis of the "W-site" and "L-site" hydrophobic pockets in GABARAP and LC3 proteins.

The presence of the hydrophobic amino acids (**W/Y/F**xx**L/I/V**) in the LIR and their interaction with the "W-site" and "L-site" hydrophobic surfaces are considered the important features of a typical LIR. Comparison of the unliganded crystal structures of ATG8 family proteins demonstrates that GABARAP and GABARAPL1 have deeper "W-site" and "L-site" hydrophobic pockets than LC3A, LC3B and LC3C (Figure 4.9), with the difference particularly pronounced for the former site. This suggests that the GABARAP family proteins have favourable architecture for accommodating LIRs with bulkier aromatic amino acids, such as tryptophan, more efficiently. This difference in the "W-site" and "L-site" hydrophobic pockets may explain the higher binding affinity of GABARAP proteins with the PLEKHM1 LIR.

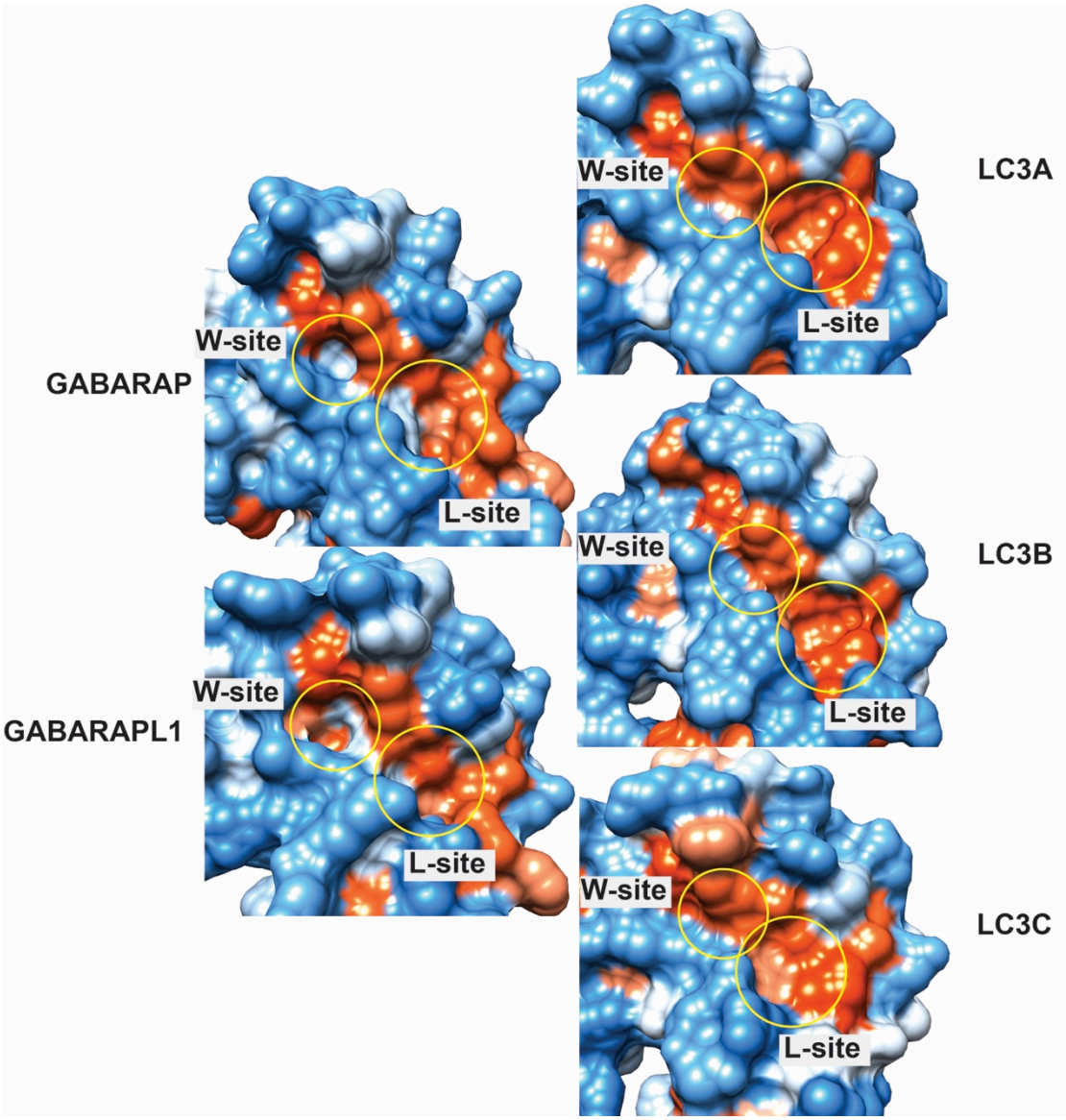


Figure 4.9: Analysis of “W-site” and “L-site” hydrophobic pockets (yellow circles) of human ATG8 family proteins. There is a noticeable difference in the depth of the “W-site” and the “L-site” hydrophobic pockets, with GABARAP proteins having deeper pockets than LC3 proteins. Surface representations of GABARAP and LC3 proteins are coloured based on hydrophobicity using the Kyte-Doolittle algorithm. The most hydrophobic and hydrophilic regions are shown in orange-red and blue, respectively. The figure was generated using Chimera.

4.3 Discussion

The PLEKHM1 LIR peptide binds the GABARAP family proteins with higher affinity than LC3 family proteins, as demonstrated by ITC experiments. In addition to the large differences in binding affinity, GABARAP family proteins demonstrated entropy values that were negative, in contrast to those of LC3A and LC3B.

The interaction of PLEKHM1 LIR with human ATG8 family members was studied by solving the crystal structures of the PLEKHM1 LIR in complex with LC3A, LC3C, GABARAP and GABARAPL1 using two approaches. In the first approach, the PLEKHM1 LIR was expressed as a chimeric fusion protein to produce the PLEKHM1 LIR-LC3A, PLEKHM1 LIR-GABARAP and PLEKHM1 LIR-GABARAPL1 complexes. In the second approach, the PLEKHM1 LIR peptide was co-crystallised with LC3C. An overlay of monomers in each asymmetric unit demonstrated similar orientation of the PLEKHM1 LIR across all the monomers, suggesting that the LIR interaction observed in the crystal structure is conserved.

Despite the similar interaction of the PLEKHM1 LIR across all five human ATG8 structures, subtle differences contribute to a tighter interaction between the PLEKHM1 LIR and GABARAP proteins, in comparison with LC3 proteins. Structural comparison of the PLEKHM1 LIR-human ATG8 complexes showed that the lysine and leucine residues located in the β_2-sheet in all five human ATG8 proteins form hydrogen bonds with the Val636 and Val638 of the PLEKHM1 LIR. However, the presence of arginine at position 28 in GABARAP, instead of the lysine residue found in the corresponding position in LC3 proteins, allows interaction with Asn637 in the PLEKHM1 LIR, similar to the interaction of Trp635 in the PLEKHM1 LIR with Glu17 in GABARAP.

Given the importance of hydrophobic surfaces in a typical LIR interaction, it was noted that the "W-site" and "L-site" hydrophobic pockets were deeper in GABARAP proteins than in LC3 proteins. This suggests that the hydrophobic amino acids Trp635 and Val638 of the PLEKHM1 LIR are buried deeper in the "W-site" and "L-site" pockets of GABARAP proteins, when compared with LC3 proteins.

In conclusion, the above discussion provides a rationale for the difference in binding affinity of the PLEKHM1 LIR with GABARAP proteins and LC3 proteins by analysing the crystal structures of PLEKHM1-human ATG8 family complexes. Comparison of amino acids in the α_3-helix shows that GABARAP proteins have more hydrophobic residues than LC3 proteins. Therefore, in the next chapter, the effects of mutating single amino acid residues in the α_3-helix on the binding affinity for the LIR will be studied.

Chapter 5. THE "W-SITE" HYDROPHOBIC POCKET REGULATES THE INTERACTION OF THE LIR IN LC3A AND GABARAP IN DIFFERENT WAYS.

5.1 Introduction

In the previous chapter, it was suggested that the interaction between the Glu17 in GABARAP and Trp635 in the PLEKHM1 LIR, and the interaction between Arg28 in GABARAP and Asn637 in the PLEKHM1 LIR, contribute to the higher binding affinity of the PLEKHM1 LIR towards GABARAP than to LC3A. It was also shown that the "W-site" hydrophobic pocket of GABARAP proteins is noticeably deeper than LC3 family proteins. Together, the results suggested that there is a difference between GABARAP proteins and LC3 proteins in their mechanism of interaction with the LIR. In this chapter, two key questions are addressed. The first is how the "W-site" hydrophobic pocket mediates the LIR interaction differently between LC3A and GABARAP family proteins. The second question is what underlying mechanism regulates the "W-site" hydrophobic pocket differently between LC3A and GABARAP family proteins.

To answer the above questions, the ULK1 LIR was used to study the differences in the LIR-dependent interactions of GABARAP and LC3A proteins. The rationale for switching from the PLEKHM1 LIR to the ULK1 LIR was to eliminate the hydrogen bonds formed by the side chains of Trp635 and Asn637 of the PLEKHM1 LIR with GABARAP proteins, but not with LC3 proteins, that contribute to the LIR interaction. The ULK1 LIR sequence (SCDTDDFVMVPA) does not make

electrostatic interactions with LC3A or GABARAP, and thus was chosen for further experiments.

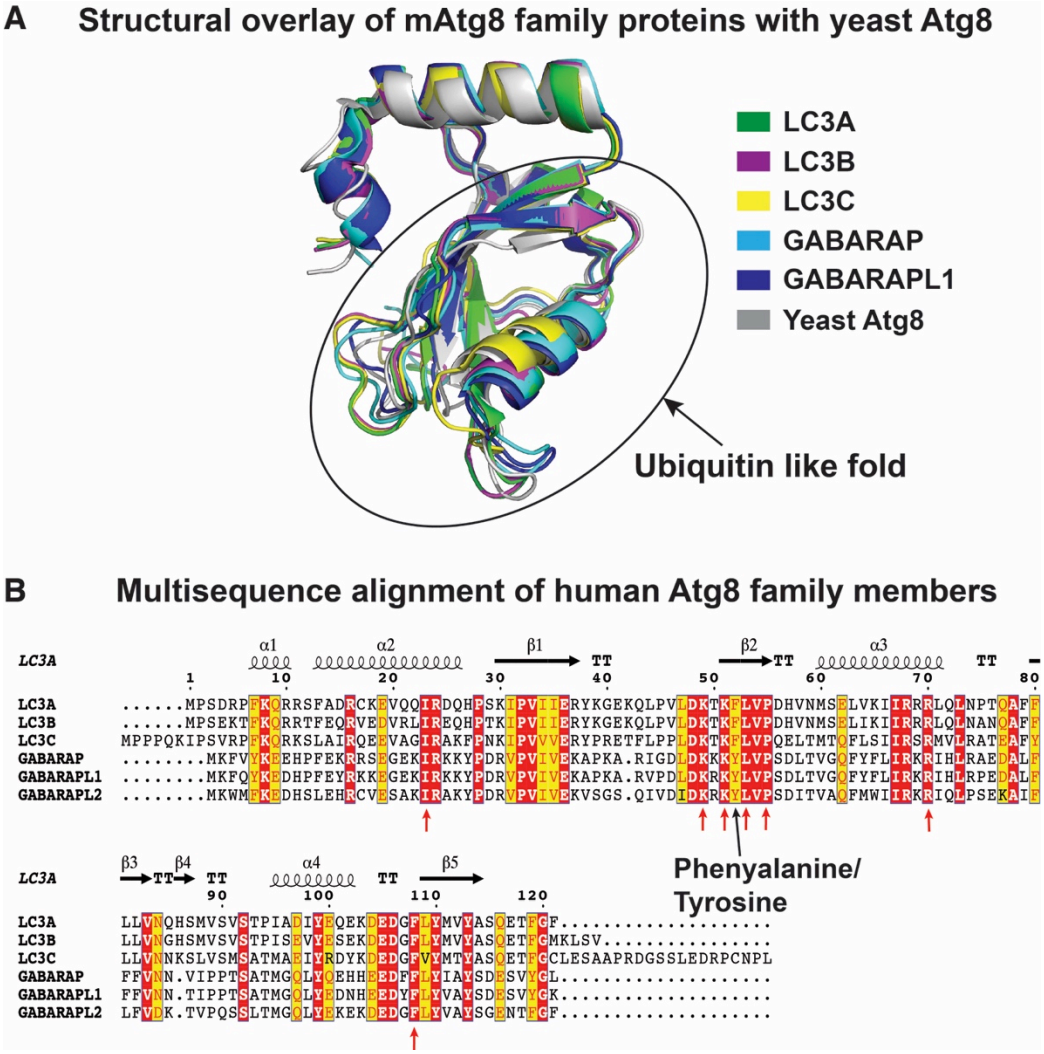


Figure 5.1: A. Structural overlay of LC3A (green), LC3B (magenta), LC3C (yellow), GABARAP (cyan), GABARAPL1 (blue) and yeast Atg8 (grey). The mammalian Atg8 family proteins have a similar ubiquitin-like fold to yeast Atg8. B. Multiple alignment of human ATG8 family proteins showing a high sequence similarity (96%) between LC3 family proteins and GABARAP family proteins. Residues marked in red are identical and residues marked in yellow have high similarity. The residues involved in LIR interactions (Atg13, p62, PLEKHM1, Optinurin, Nix and ALFY) are shown by a red arrow, and all other interacting residues are identical, except for Phe52 in LC3 proteins and Tyr49 in GABARAP proteins. The sequence alignment was performed using T-coffee server and annotated using the ESPript server.

Mammalian Atg8 homologues have very similar structures, and interact with their binding partners in a similar fashion (Sugawara *et al.*, 2004). They have a core ubiquitin-like fold, and their overall secondary and tertiary structures are similar to yeast Atg8 (Figure 5.1A) (Paz *et al.*, 2000, Bavro *et al.*, 2002, Sugawara *et al.*, 2004). Not surprisingly, the human ATG8 homologues share high sequence similarity (96%) (Figure 5.1B). Despite their similar structures and sequences, the GABARAP and LC3 proteins interact with the LIRs with different binding affinities. This suggests that the mode of interaction with the LIR differs between GABARAP and LC3 proteins.

In this chapter, a combination of surface plasmon resonance (SPR), X-ray crystallography and molecular dynamics simulations were used to demonstrate that mutation of a single amino acid affects the "W-site" hydrophobic pocket of LC3A, reducing its binding affinity with the ULK1 LIR, whereas the corresponding mutation in GABARAP does not affect the binding affinity with the ULK1 LIR. This difference in binding mechanism at the "W-site" hydrophobic pocket suggests that this single amino acid controls the ULK1 LIR-"W-site" interaction in LC3A, but not in GABARAP.

5.2 Results

5.2.1 Purification of LC3A and GABARAP.

To study the crystal structures and perform SPR experiments, pure protein samples were required. To achieve this, DNA encoding LC3A (2–121) and GABARAP (2–117) was cloned into the pET30ΔSE vector, which expresses with a His_6 tag. The modified pET30ΔSE vector, which has the serine and glutamic acid residues deleted from the linker region between the tag and cloning sites of pET30, was used to reduce the number of residues in the tagged protein that are not part of the protein of interest. Protein was expressed from these vectors in *E. coli* BL21(DE3). The His_6-tagged LC3A and GABARAP proteins were isolated and purified from the cell lysate using Ni-NTA affinity chromatography. Following the removal of the His_6 tag by thrombin cleavage, the protein was purified by SEC, as shown by sharp peaks in the chromatograms (Figure 5.2A and 5.2B). Samples from the peak fractions were analysed on an SDS-PAGE gel (Figure 5.2C), which showed bands at ~15 kDa, corresponding to the molecular weights of LC3A (2–121) and GABARAP (2–117).

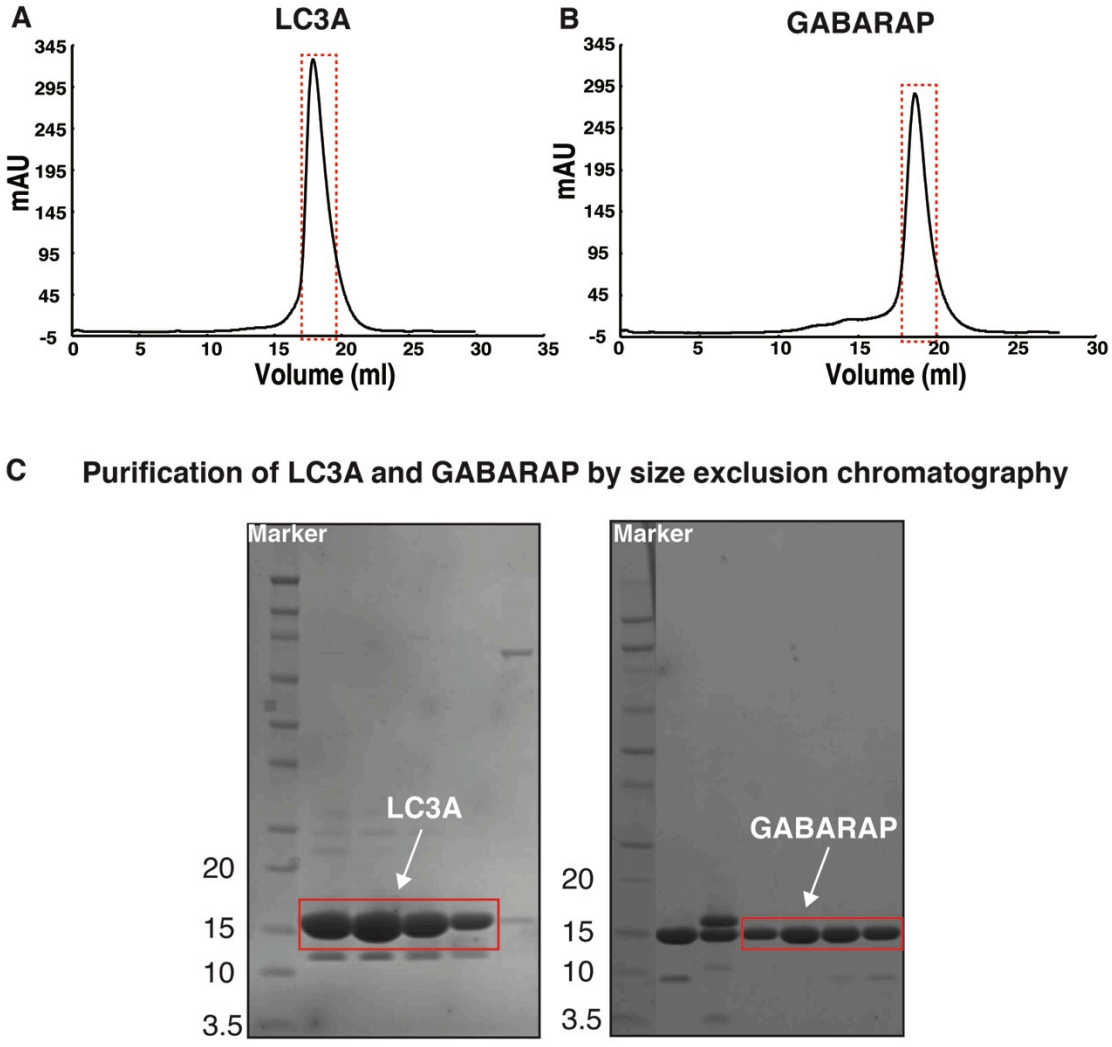


Figure 5.2: A&B: SEC chromatograms for the purification of LC3A and GABARAP, respectively. The purified proteins were visualised as a distinct peak in each of the chromatograms (red box). The samples from these fractions were analysed on an SDS-PAGE gel (C), which shows bands for the purified LC3A and GABARAP at the expected size (~15 kDa).

5.2.2 GABARAP has a higher binding affinity for the ULK1 LIR peptide than LC3A.

Using SPR fitted to the 1:1 Langmuir isotherm model (Figure 5.3A–D), the binding affinity of LC3A and GABARAP with the ULK1 LIR peptide (KSCDTDDFVMVPAQ) was measured. To determine the association and dissociation constants, LC3A and GABARAP were injected as analytes, and the ULK1 LIR peptide was immobilised to the chip. GABARAP bound the ULK1 LIR

peptide with significantly higher affinity (K_D = 0.3 ± 0.05 μM) than the LC3A protein (K_D = 4.4 ± 0.5 μM). A scrambled version of the ULK1 LIR peptide (KVSFPQACDMDTVD) was used as a negative control, and did not bind to either LC3A or GABARAP. The data was consistent across the triplicate samples collected for both proteins (Figure 5.4). The k_{on} and k_{off} rates were very fast for both proteins, such that it was difficult to fit these with meaningful values (Figure 5.4).

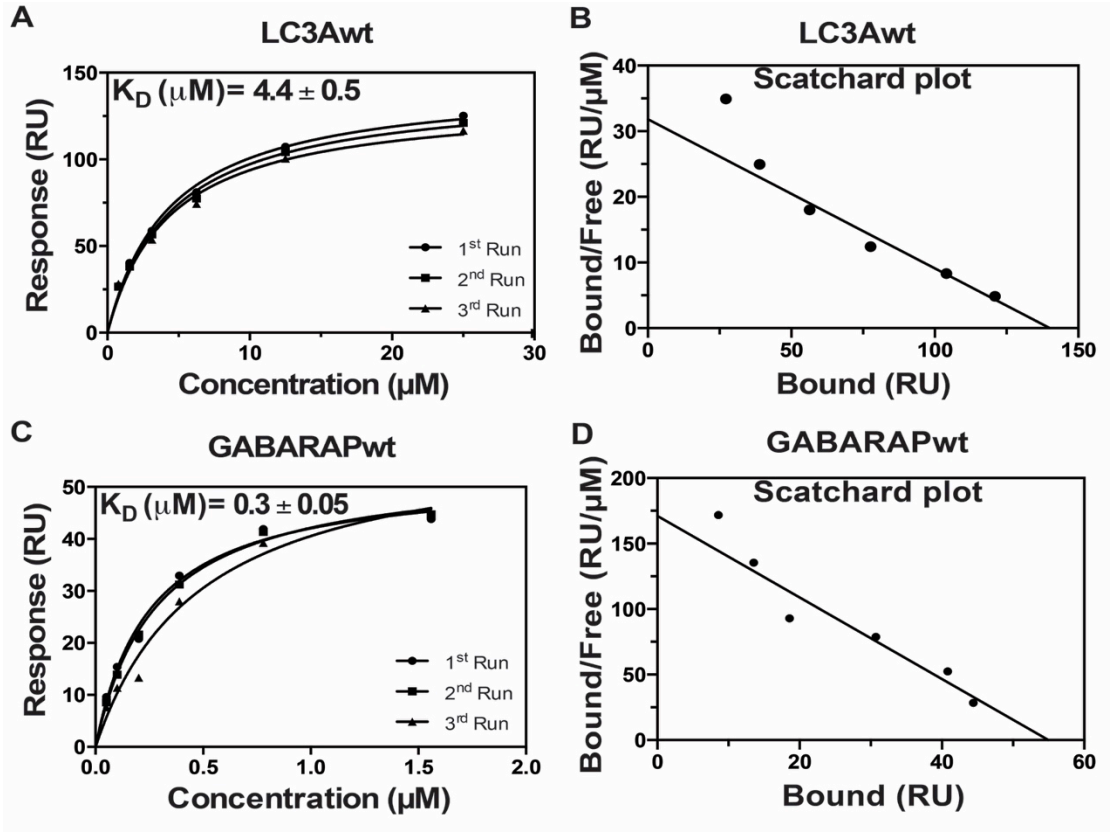


Figure 5.3: A & C. 1:1 Langmuir model for wild-type (wt) LC3A and GABARAP fitted for the triplicate data sets. The equilibrium-binding constant was calculated from the response vs. concentration curve. The K_D value of GABARAP (K_D = 0.3 ± 0.05 μM) showed it had a 12-fold higher binding affinity than LC3A (K_D = 4.4 ± 0.6 μM) for the ULK1 LIR. B & D. Scatchard plot for the average from the triplicate data showed the relative binding affinity and the difference in binding affinity between LC3A and GABARAP. GABARAP had higher binding affinity and agreed with the 1:1 Langmuir model.

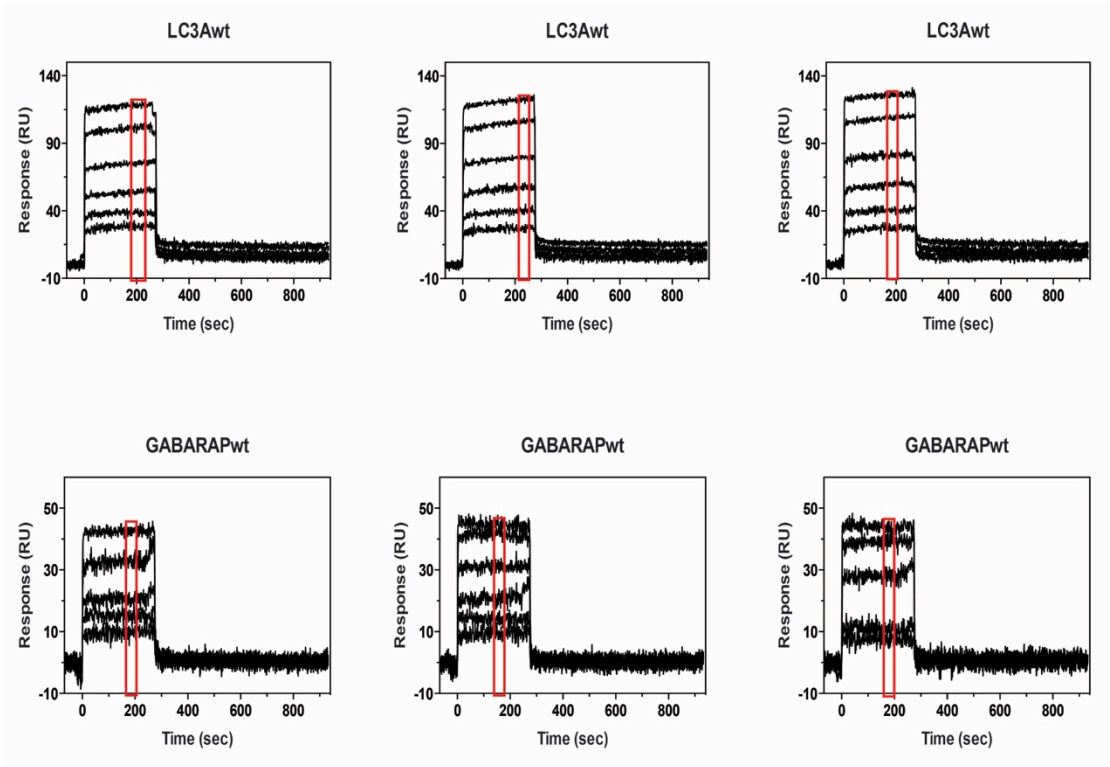


Figure 5.4: Triplicate SPR response data collected for wt LC3A and GABARAP. The response data is shown in black, and six concentrations of LC3A (25 μM, 12.5 μM, 6.25 μM, 3.125 μM, 1.56 μM and 0.78 μM) and six concentrations of GABARAP (1.56 μM, 0.78 μM, 0.39 μM, 0.195 μM, 0.097 μM and 0.048 μM) were injected. The data points from the region marked by the red box were averaged and used in the 1:1 Langmuir model to calculate the binding affinity.

5.2.3 Crystal structure of the ULK1 LIR fused to GABARAP demonstrates that LIR interaction is similar in LC3A and GABARAP

Given the close sequence and structural similarities of the unliganded GABARAP and LC3A proteins (Figure 5.1B), it was surprising that these proteins bind the ULK1 LIR peptide with different affinities. To address this at the molecular level, the crystal structure of GABARAP in complex with the ULK1 LIR was solved (to 2.09-Å resolution, data collection and refinement statistics in **Table 5.1**), and compared with the previously reported LC3A-ATG13 LIR structure ((Suzuki *et al.*, 2014), PDB code: 3WAN; ATG13 LIR sequence DDFVMI). The ULK1 LIR peptide (354–TDDFVMV–360) was fused at the N-terminus of GABARAP, generating a GABARAP-ULK1 LIR chimera. In the structure, the ULK1 LIR peptide interacted

with a symmetry-related GABARAP molecule in the crystal lattice (Figure 5.5A). The aromatic amino acid Phe357 of the ULK1 LIR docked into the "W-site" hydrophobic pocket, and Val360 interacted at the "L-site" hydrophobic pocket (Figure 5.5A). The main chains of Val358 and Val360 formed hydrogen bonds with the main chains of Leu50 and Lys48 in GABARAP. The main chain oxygen of Val360 also formed a hydrogen bond with side chain nitrogen of Arg28 at the (NH1) position (Figure 5.5B).

Table 5.1: Summary of data collection and refinement of the ULK1 LIR (354–360) in complex with GABARAP.

	ULK1$^{354–360}$-GABARAP
Data collection	
Space group	$P3_221$
Cell dimensions (Å)	a = 58.2, b = 58.2, c = 83.5 α = 90, β = 90, γ = 120
Wavelength (Å)	0.9537
Resolution (Å)	43.15-2.09 (2.15-2.09)
R_{merge}	0.082 (0.678)
$R_{p.i.m.}$	0.086 (0.713)
$R_{meas.}$	0.205 (3.652)
$I/\sigma I$	10.3 (1.5)
Completeness	99.8 (98.0)
CC(1/2)	0.998 (0.897)
Redundancy	10.6 (10.2)
Refinement	
Total reflections (unique)	106,643 (8,234)
R_{work}/R_{free}	0.206/0.245
Number of atoms	
Protein	1,025
Water	44
B-Factors	
Main chain	34.4
Side chain	37.1
LIR	39.1
Solvent	37.4
Ramachandran plot	
Allowed regions	100
Outliers	-
R.m.s.d. from ideality	
Bond lengths (Å)	0.007
Planarity (°)	0.004
Angles (°)	1.033

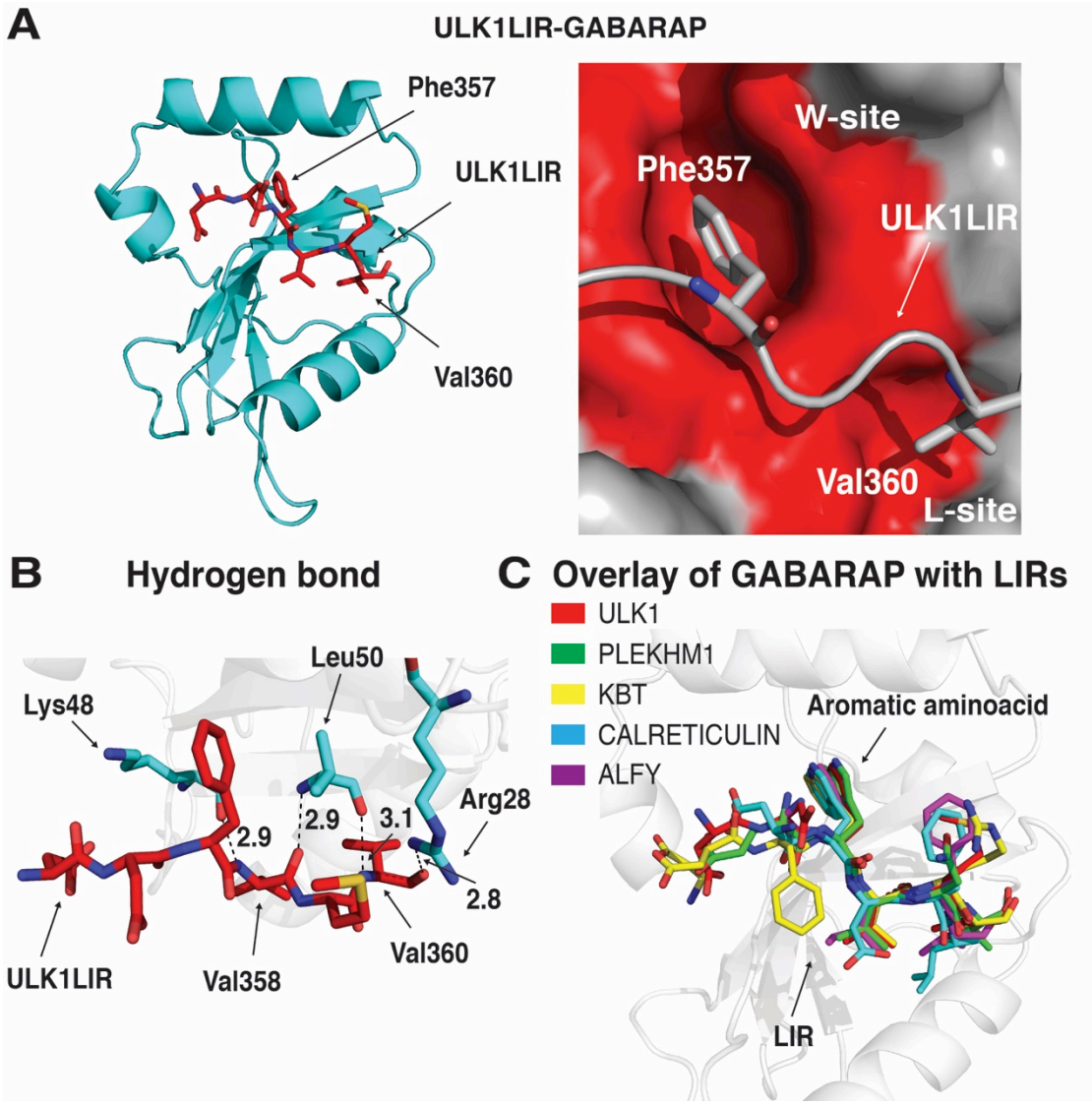


Figure 5.5: A. (Left panel) Cartoon representation of ULK1 LIR-GABARAP (cyan) showing the interaction of the ULK1 LIR (red sticks) with GABARAP, and (right panel) surface representation of "W-site" and "L-site" hydrophobic pockets (in red) of GABARAP showing the position at which Phe357 and Val360 of the ULK1 LIR interact with the hydrophobic pocket. B. Backbone of Val358 and Val360 in ULK1 LIR (red) forms hydrogen bonds with residues from GABARAP (cyan). C. Structural overlay of various LIRs interacting with GABARAP. All the LIRs examined interact with the ULK1 LIR in the same conformation.

The binding conformation of the peptide against the GABARAP surface is similar to that of other ATG8-LIR complexes (Figure 5.5C) (Genau *et al.*, 2015, Thielmann *et al.*, 2009, Lystad *et al.*, 2014). Structural comparison between LC3A-ATG13 LIR ((Suzuki *et al.*, 2014), PDB code 3WAN) and GABARAP-ULK1 LIR shows a similar orientation of the LIR peptide in both the LC3A and GABARAP proteins (Figure 5.6A).

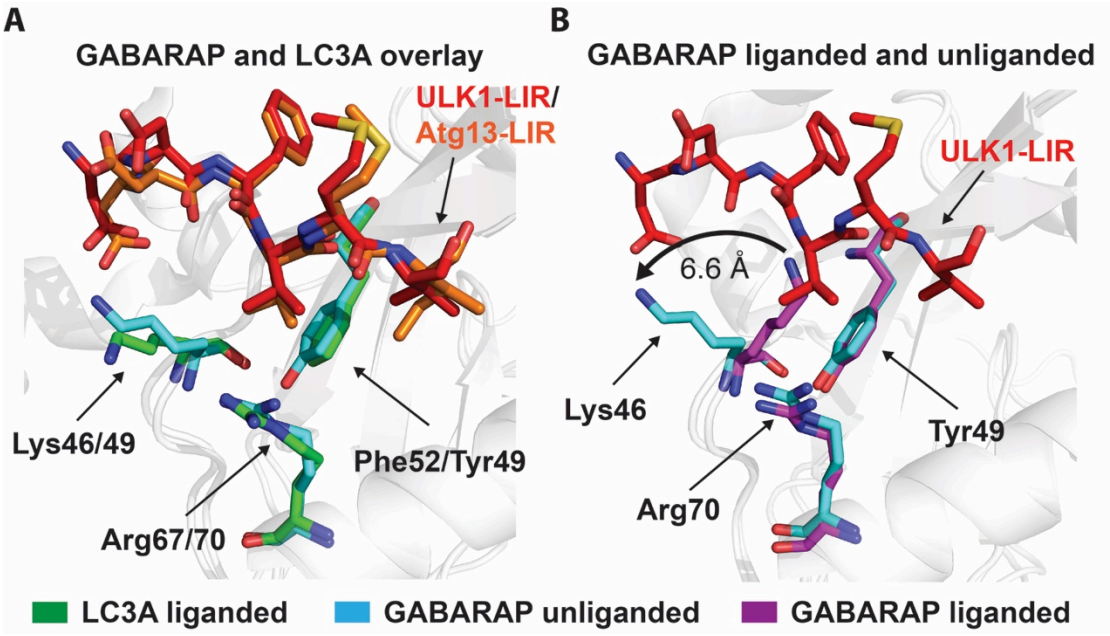


Figure 5.6: A. Structural comparison of ATG13 LIR-LC3A residues (green) and ULK1 LIR-GABARAP residues (cyan) with the ATG13 LIR peptide in orange, and the ULK1 LIR peptide in red. This demonstrated the presence of a tyrosine residue in GABARAP instead of phenylalanine in LC3A at the corresponding position. B. Structural comparison of GABARAP in liganded and unliganded forms. The residues in purple were in complex form, and the residues in cyan were in unliganded form. A 6.6-Å structural rearrangement of the Lys46 side chain was observed between the two forms.

Structural comparison of the unliganded (PDB code; 1GNU) and the liganded forms of GABARAP (Figure 5.6B) demonstrated that Lys46 gates peptide binding at the "W-site" site, as recently demonstrated for LC3X (Suzuki *et al.*, 2014). In this mechanism, the amino group of Lys46 moved 6.6 Å between its open and closed states. The most obvious sequence difference between LC3A and GABARAP at the "W-site" is the presence of Tyr49 in GABARAP, instead of Phe52 in LC3A (Figure 5.6B). The sequence alignment between all human ATG8 proteins (Figure 5.1B) demonstrates that the residues involved in LIR interaction (marked with red arrows) were conserved, except that Tyr and Phe were exchanged.

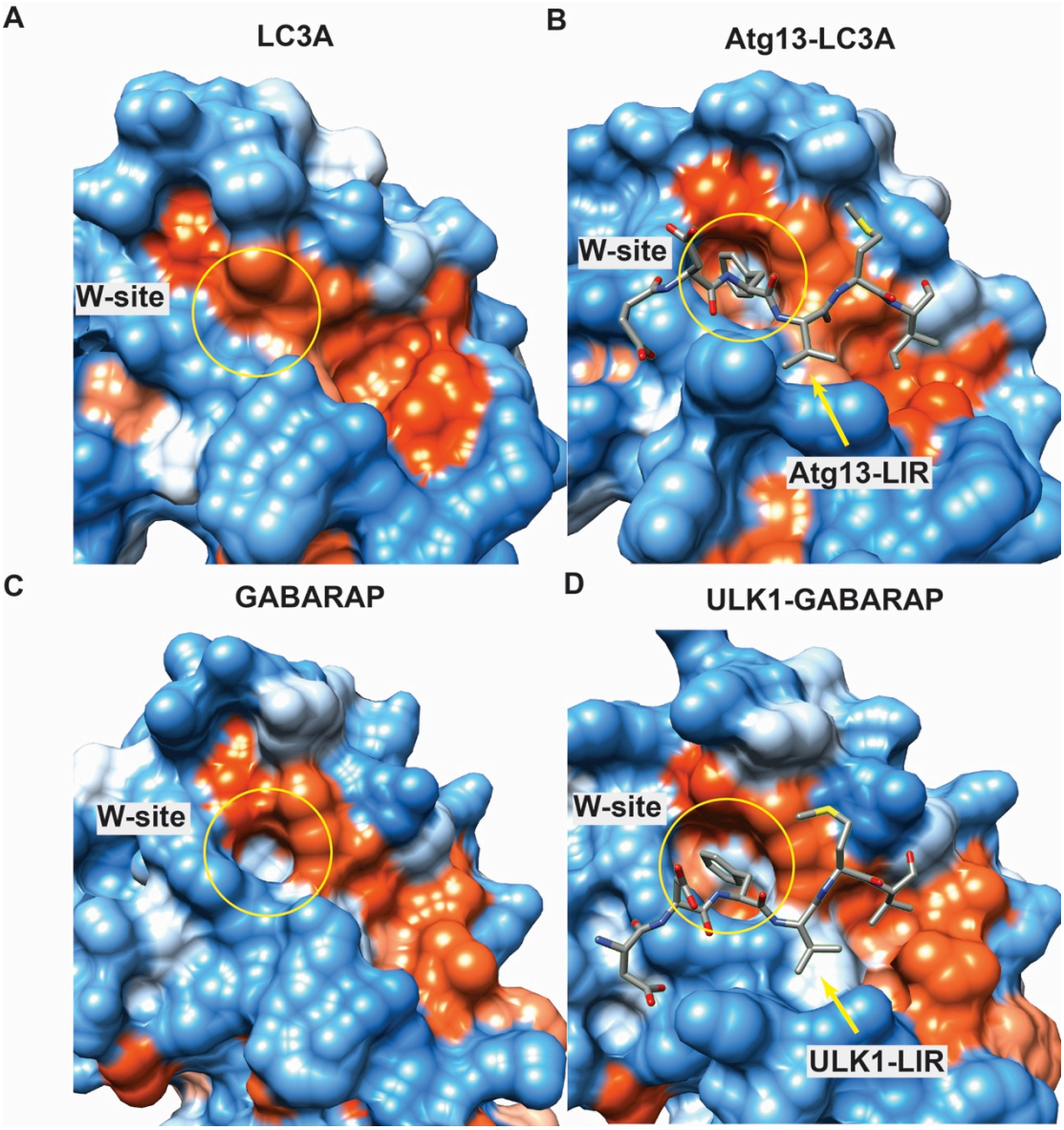


Figure 5.7: A & B show a surface representation of LC3A in unliganded (A) and liganded forms (B) with the ATG13 LIR. The most hydrophobic and hydrophilic regions, as determined by the Kyte-Doolittle algorithm, are shown in orange-red and blue, respectively. The figure was generated using Chimera. The "W-site" hydrophobic pocket, marked with a yellow circle, shows a closed conformation in the unliganded form and an open conformation in the liganded form. C & D show a surface representation of GABARAP in unliganded (A) and liganded forms (B) with the ULK1 LIR. Colouring is as described in panels A and B. The "W-site" hydrophobic pocket, marked with a yellow circle, shows an open conformation in both unliganded and liganded forms.

Structural comparison of the unliganded (GABARAP PDB code: 1GNU; LC3A PDB code: 3WAL) and liganded (LC3A PDB code: 3WAN) forms of GABARAP and LC3A using surface representation (Figure 5.7) showed that the "W-site" hydrophobic pocket in LC3A opens only in the liganded state, and remains

closed in the unliganded state (Figure 5.7A and B). However, in GABARAP, the "W-site" hydrophobic pocket remains open regardless of the ligand-binding state (Figure 5.7C and D). This clearly shows that the difference in the LIR-binding mechanisms of LC3A and GABARAP occurs at the "W-site" hydrophobic surface.

5.2.4 Molecular dynamics simulations demonstrate that Arg70 in LC3A and Arg67 in GABARAP interact differently with the ULK1 LIR.

Given that there is a noticeable structural difference at the "W-site" hydrophobic surface between liganded and unliganded structures of GABARAP and LC3A, the dynamics of the structures were investigated to verify: 1) whether LC3A is more dynamic than GABARAP; 2), whether Phe357 and Val360 of the ULK1 LIR interact differently with the "W-site" and the "L-site" of LC3A and GABARAP; and 3) whether there are any other amino acids that play key roles in the interaction with the ULK1 LIR. To investigate these questions, a 100-ns molecular dynamics simulation was performed on the crystal structures of unliganded and liganded LC3A and GABARAP. The LIR sequence of ATG13 (FVMI) was modified to the ULK1LIR sequence (FVMV) in the ATG13 LIR-LC3A (3WAN) crystal structure and used as the initial coordinate file for the ULK1 LIR-LC3A simulation.

The RMSD calculated for the α-carbon with the starting frame as the reference (Figure 5.8) showed that the LC3A structures were more dynamic than the GABARAP structures in both liganded and unliganded simulations. The RMSD of ULK1 LIR-GABARAP displayed a more stable simulation than unliganded GABARAP, suggesting that interaction with the ULK1 LIR may stabilise GABARAP (Figure 5.8A and D). However, no such differences were noticed in the RMSD of liganded and unliganded LC3A simulations (Figure 5.8B and E). Overall, the RMSD

also provided information about the quality of the simulation, and suggested that all six simulations were stable over a period of 100 ns (Figure 5.8A–F). To support the simulation data that suggested LC3A was more dynamic, thermostability assays were performed using differential scanning fluorimetry (DSF) (Figure 5.8G). The results correlated with the simulation data, as the less dynamic protein, GABARAP, had a higher melting point (55°C) than LC3A (47°C) (Figure 5.8G), indicating that it was more stable.

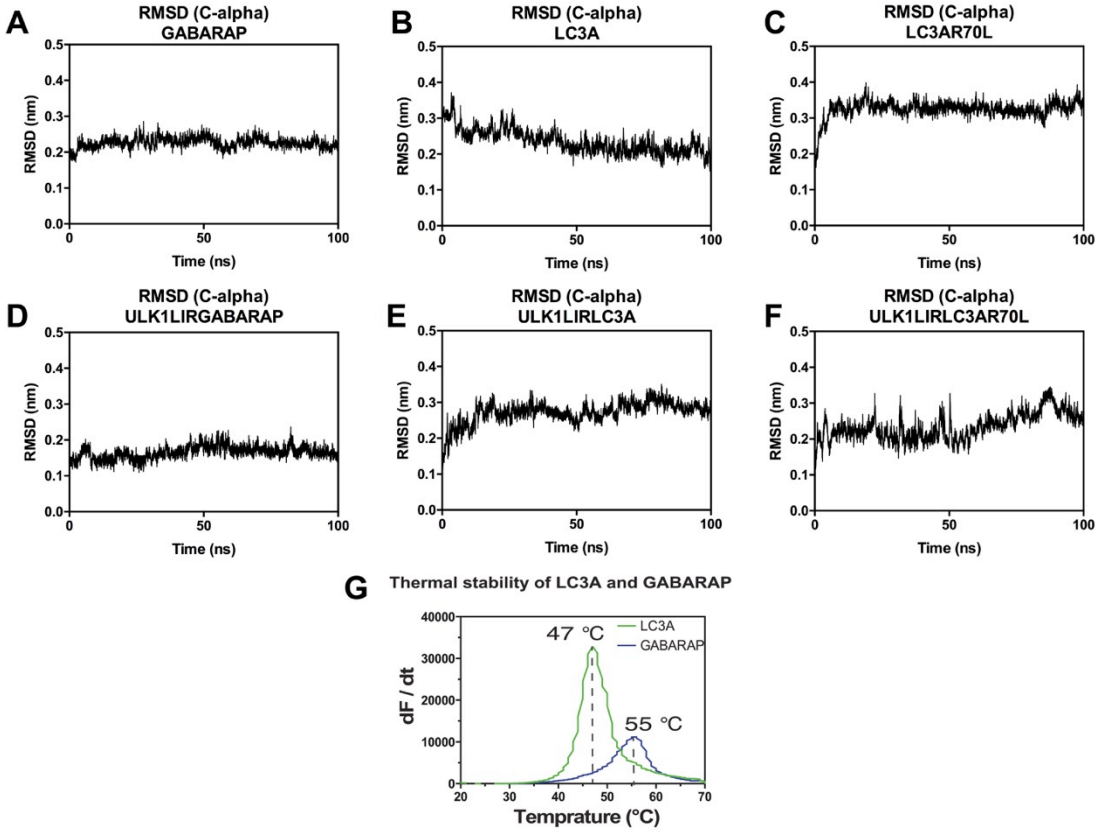


Figure 5.8: Panels A–F show the RMSD of GABARAP, LC3A, LC3A-R70L, ULK1 LIR-GABARAP, ULK1 LIR-LC3A and ULK1 LIR-LC3A-R70L, respectively. The RMSD was calculated for the α-carbon over the entire duration of the simulation (100 ns) using the gmx rms function of GROMACS and the starting frame as the reference frame. The plot demonstrates a more dynamic simulation for LC3A and ULK1 LIR-LC3A (around 0.3 nm) than for GABARAP and ULK1 LIR-GABARAP (≤0.2 nm). G. The DSF plot shows GABARAP has greater thermal stability (melting temperature: 55°C) than LC3A (melting temperature: 46°C).

Next, the "W-site" and "L-site" hydrophobic pockets of LC3A and GABARAP were analysed using the simulation data. The crystal structure clearly

showed that the LC3A "W-site" hydrophobic pocket was more dynamic between its LIR-bound and -unbound states than GABARAP. Therefore, it was interesting to analyse the RMSD of ULK1 LIR amino acids Phe357, which interacts at the "W-site", and Val360, which interacts at the "L-site" (Figure 5.9). The RMSD for α-carbons was calculated by aligning the ULK1 LIR peptide with the protein (LC3A or GABARAP). Consistent with the SPR data, which showed LC3A had lower binding affinity for the ULK1 LIR than GABARAP, LC3A had a more dynamic Phe357 at the "W-site" than GABARAP (Figure 5.9A). However, Val360 at the "L-site" did not show much dynamic difference between LC3A and GABARAP (Figure 5.9B). This result, consistent with both SPR data and structural comparison, clearly indicated that the weaker binding affinity of LC3A towards the ULK1 LIR is caused by the dynamic "W-site" hydrophobic pocket of LC3A, which does not make a strong hydrophobic interaction with Phe357 of the ULK1 LIR.

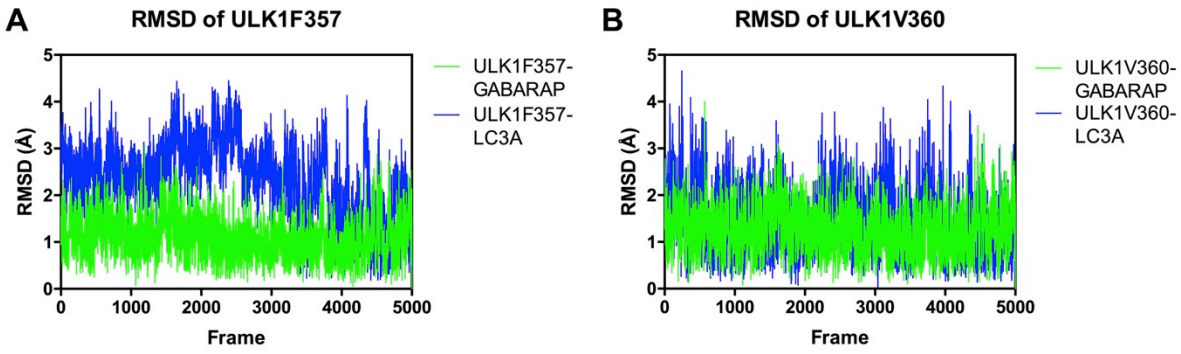


Figure 5.9: A. Comparison of the RMSD of ULK1 LIR Phe357, which interacts with the "W-site" in both LC3A (blue) and GABARAP (green), shows greater flexibility in LC3A, suggesting a weaker interaction of ULK1 LIR with LC3A compared with that of GABARAP. B. Comparison of RMSD of ULK1 LIR Val360, which interacts at the "L-site" in both LC3A (blue) and GABARAP (green), does not show much difference in the flexibility between the two proteins.

Finally, the simulation data were used to identify the amino acids in LC3A and GABARAP that interacted differently with the ULK1 LIR. Considering the lower binding affinity of LC3A towards the ULK1 LIR, we were interested to determine

whether there were any additional interactions that supplemented the hydrophobic interaction in LC3A. A visual inspection of the frames from the simulation of ULK1 LIR-LC3A and ULK1 LIR-GABARAP showed that the guanidinium side chain of Arg70 in LC3A was in close proximity to the main chain oxygen of Met359 of the ULK1 LIR. Consequently, the distance between the guanidinium side chain of Arg70 and the main chain oxygen of Met359, and the corresponding residues in GABARAP (side chain guanidinium of Arg67 and main chain oxygen of Met359), was calculated (Figure 5.10). Interestingly, the side chain of LC3A Arg70 made a consistent electrostatic interaction throughout the simulation at a distance of less than 3 Å (Figure 5.10B). To verify that it was an interaction between Arg70 and Met359, and not an artefact resulting from the dynamic nature of LC3A, Arg70 was replaced with leucine, which does not have a positively charged side chain. As expected, the Leu70 side chain was 6–8 Å from Met359, which is beyond the distance for electrostatic interaction (Figure 5.10C). The side chain of Arg67 in GABARAP was also ~6 Å from the main chain of Met539 (Figure 5.10A). The effect of Arg70 in LC3A and Arg67 in GABARAP on the ULK1 LIR interaction was examined further using SPR.

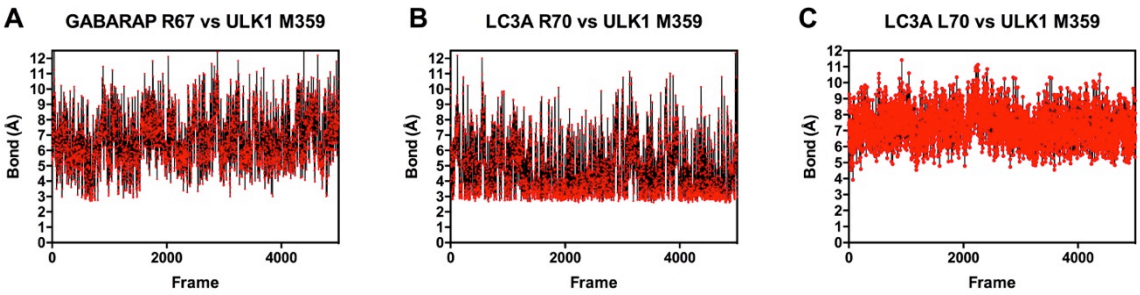


Figure 5.10: A. The distance between the side chain of Arg67 in GABARAP and Met359 in the ULK1 LIR was measured using visual molecular dynamics (shown in red dots for each frame). The average distance throughout the simulation was ~6 Å, suggesting no interaction between the two residues. B. The distance between the side chain of Arg70 in LC3A and Met359 in the ULK1 LIR was measured, and showed an average distance of <3 Å throughout

the simulation, suggesting an electrostatic interaction between the residues. C. The distance between the side chain of Leu70 in LC3A-R70L and Met359 in the ULK1 LIR was measured, and the distance of approximately 6–8 Å suggested no interaction between the residues.

5.2.5 Purification of LC3A and GABARAP wild-type protein.

Structural studies and molecular dynamics simulations suggested that two amino acid positions (Arg70 and Phe52 in LC3A and Arg67 and Tyr49 in GABARAP) may be important for the variable binding affinity between GABARAP and LC3A to the ULK LIR. To study the effect of the charge of Arg70 in LC3A and Arg67 in GABARAP in the "L-site" binding pocket, the arginine residues were mutated to leucine. To determine whether the tyrosine at position 49 in the "L-site" binding pocket in GABARAP proteins and the phenylalanine at the equivalent position in LC3 proteins was functionally significant, the two residues were swapped.

Four mutant proteins were generated: GABARAP-R67L, GABARAP-Y49F, LC3A-R70L and LC3A-F52Y. DNA encoding the four mutant proteins was cloned into pET30ΔSE, which was used to express the proteins with N-terminal His_6 tags in *E. coli* BL21(DE3). Similar to wild-type purification, the four mutant proteins were isolated and purified from the cell lysate using Ni-NTA affinity chromatography. The target protein samples were purified by SEC, following cleavage of the His_6 tag by thrombin. The purified protein samples were observed as sharp peaks in the SEC chromatograms (Figure 5.11A–D), and samples from the peak fractions were analysed by SDS-PAGE (Figure 5.11E), which revealed bands of ~15 kDa, corresponding to the expected molecular weights of the mutant proteins.

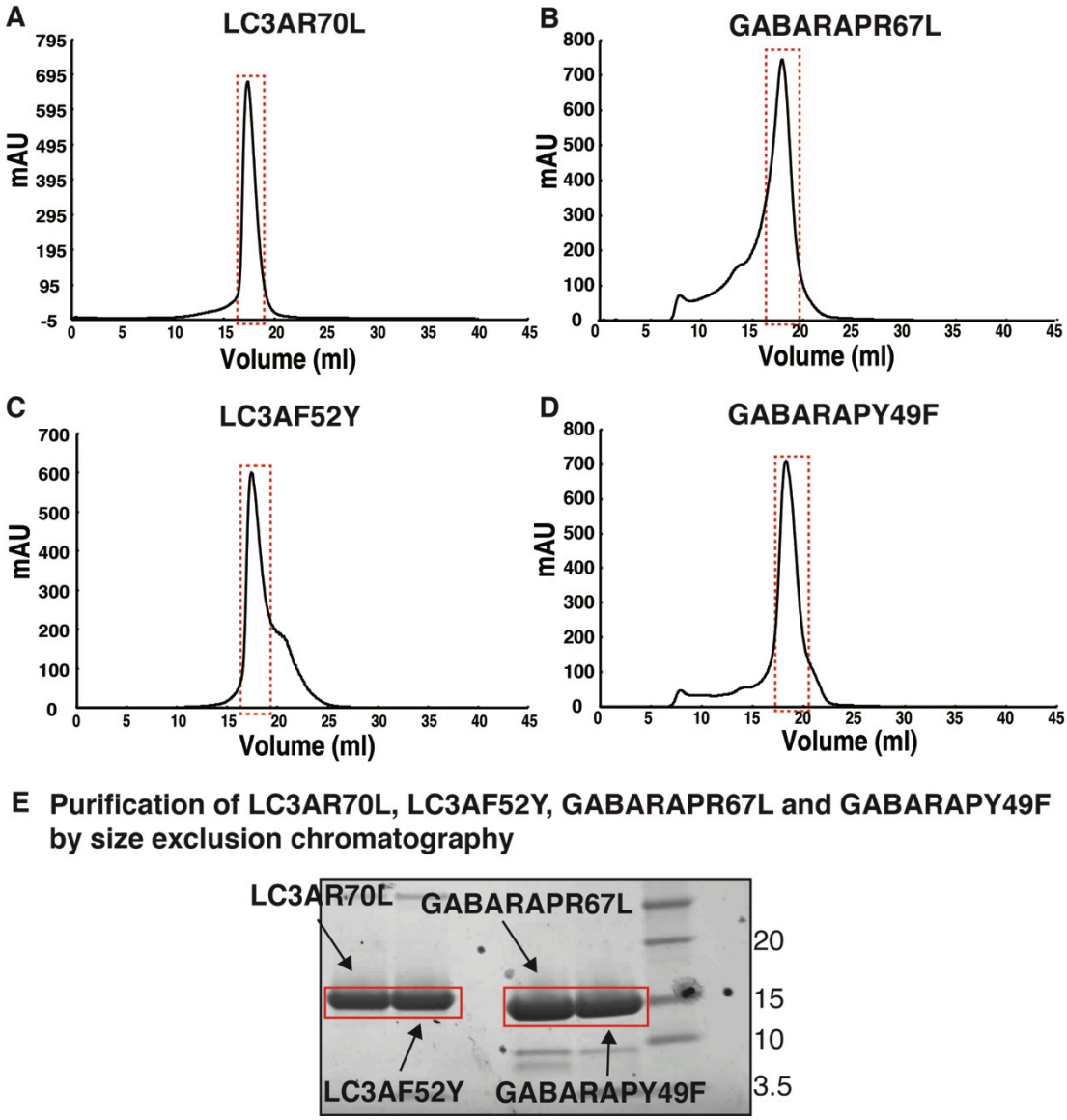


Figure 5.11: A–D. Chromatograms for SEC purification of LC3A-R70L, GABARAP-R67L, LC3A-F52Y and GABARAP-Y49F, respectively. The purified proteins were observed as distinct peaks in the chromatogram (red boxes). (E) Samples from the fractions corresponding to the peak were analysed by SDS-PAGE, which showed bands for LC3A-R70L, LC3A-F52Y, GABARAP-R67L and GABARAP-Y49F at the expected size of ~15 kDa.

5.2.6 Differences in the interactions between Arg70 in LC3A and Arg67 in GABARAP contribute to the binding affinity.

5.2.6.1 Initial binding analysis using pull-down assay shows the ULK1 LIR peptide binds to LC3A and GABARAP and their mutants.

Pull-down assays were performed to assess differences in ULK1 LIR binding affinities for LC3A, LC3A-R70L, LC3A-K49A, GABARAP, GABARAP-R67L and GABARAP-K46A (Figure 5.12). The LC3A-K49A and GABARAP-K46A mutants were used as controls because the LC3A-K49A mutant demonstrated higher binding affinity than LC3A in a pull-down assay (Suzuki *et al.*, 2014). The mutants LC3A-R70L and GABARAP-R67L bound with the GST-tagged ULK1 LIR peptide. It also confirmed the binding of LC3A and GABARAP, as reported previously. Incubation of mutants and wild-type proteins with the GST tag only demonstrated they did not interact with the GST tag. Although the pull-down assay results verified the binding of mutants and wild-type proteins with the ULK1 LIR, no differences in binding affinity were observed in spite of optimisation of the protein concentration, wash conditions and incubation time, and repeating the experiment several times.

Consequently, the more sensitive technique of SPR was used to measure the differences in binding affinity.

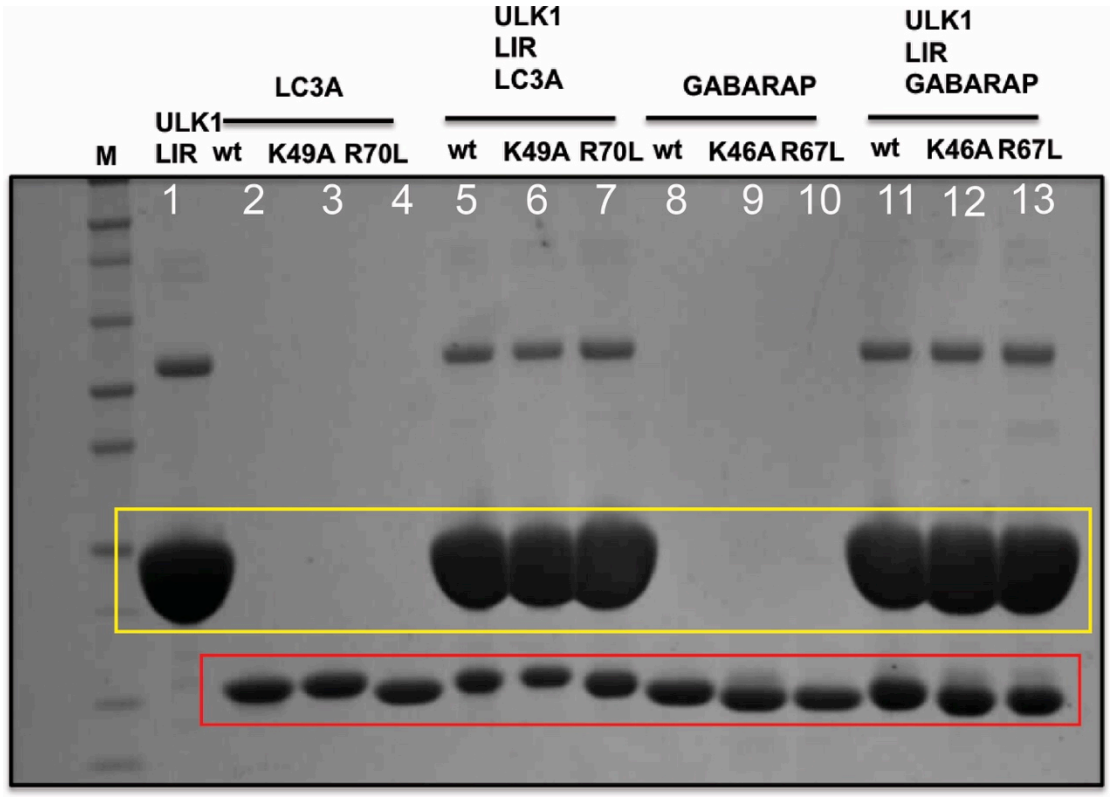


Figure 5.12: SDS-PAGE gel showing the results of the pull-down assay performed between GST-ULK1 LIR and LC3A, LC3A-K49A, LC3A-R70L, GABARAP, GABARAP-K46A and GABARAP-R67L. Lane 1: GST-fused ULK1 LIR (correct band size marked in yellow). Lanes 2–4: control proteins LC3A, LC3A-K49A and LC3A-R70L (correct band size marked in red). Lanes 5–7: pull-downs of LC3A, LC3A-K49A and LC3A-R70L with GST-fused ULK1 LIR. Lanes 8–10: control proteins GABARAP, GABARAP-K46A and GABARAP-R67L (correct band size marked in red). Lanes 11–13: pull-downs of GABARAP, GABARAP-K46A and GABARAP-R67L with GST-fused ULK1 LIR.

5.2.6.2 Differences in the interaction between Arg70 in LC3A and Arg67 in GABARAP contribute to the binding affinity.

The binding affinities of LC3A-R70L and GABARAP-R67L to the ULK1 LIR peptide were measured and compared with the wild-type protein using SPR. The dissociation constant for LC3A-R70L was 11 ± 1 μM, which causes a significant decrease in binding affinity compared with the wild-type LC3A (K_D = 4.4 ± 0.5 μM). The dissociation constant for GABARAP-R67L of 0.4 ± 0.03 μM was similar to that of the wild-type GABARAP protein (K_D = 0.3 ± 0.05 μM). The sensograms for LC3A-R70L and GABARAP-R67L followed the same trend as their wild-type counterparts: fast on and off rates (Figure 5.13A and 5.13C, replicate data shown in

Figure 5.14). Thus, an arginine residue at the "L-site" hydrophobic pocket is important for binding the ULK1 LIR in LC3A, but not for GABARAP.

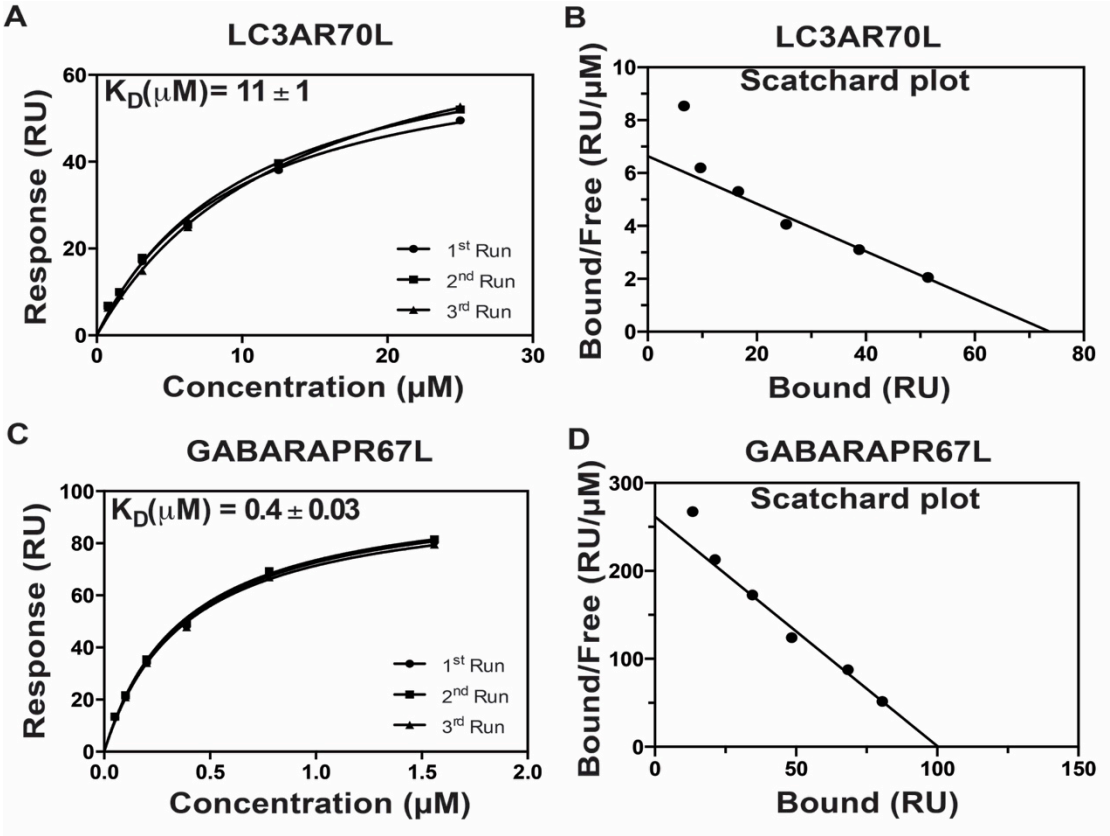


Figure 5.13: A & C. 1:1 Langmuir models for LC3A-R70L and GABARAP-R67L fitted for triplicate data set. Equilibrium binding constants were calculated from the response vs. concentration curve. The K_D value of LC3A-R70L (11 ± 1 μM) showed it had 2.5 times less binding affinity than LC3A (4.4 ± 0.6 μM). There was no difference in K_D between GABARAP-R67L and GABARAP. B & D. Scatchard plots for the average of the triplicate data shows GABARAP-R67L has higher binding affinity than LC3A-R70L, which agreed with the 1:1 Langmuir model.

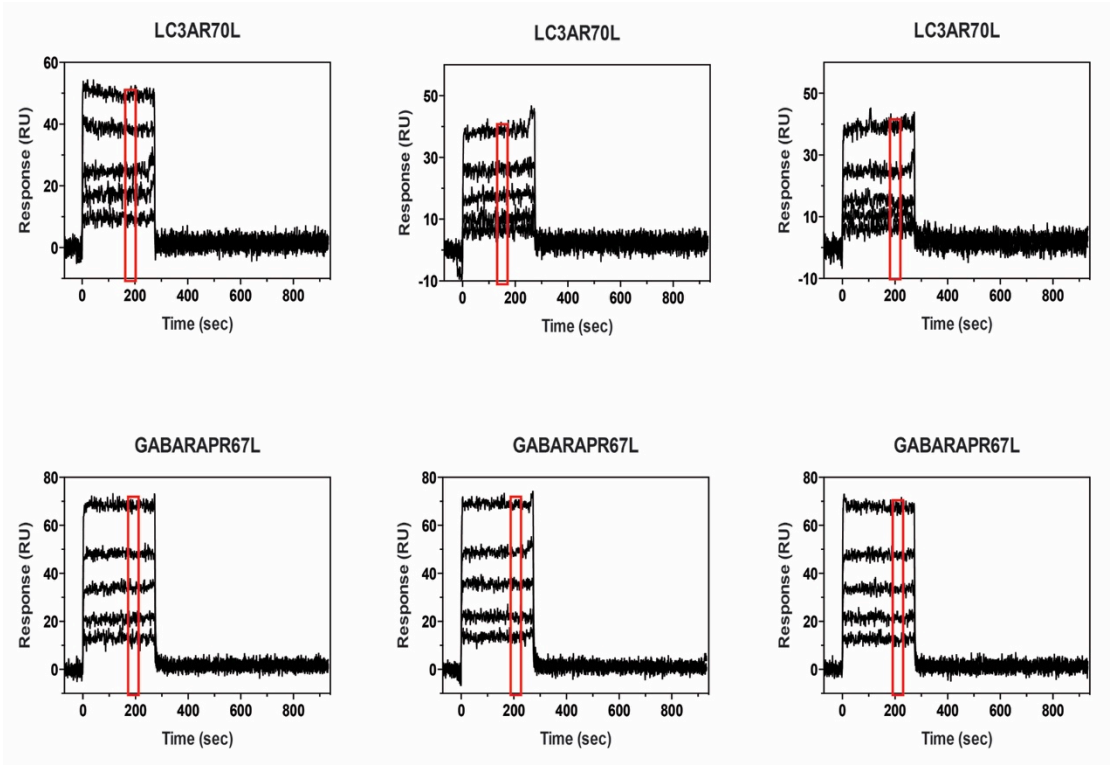


Figure 5.14: SPR response data collected in triplicate for LC3A-R70L and GABARAP-R67L. The response data is shown in black, and six concentrations of both LC3A-R70L (25 μM, 12.5 μM, 6.25 μM, 3.125 μM, 1.56 μM and 0.78 μM) and GABARAP-R67L (1.56 μM, 0.78 μM, 0.39 μM, 0.195 μM, 0.097 μM and 0.048 μM) were injected. The data points from the regions marked by red boxes were used in the 1:1 Langmuir model to calculate the binding affinity.

5.2.6.3 Presence of Tyr49 in GABARAP or Phe52 in LC3A has little effect on binding affinity for the ULK1 LIR peptide.

The dissociation constants of LC3A-F52Y and GABARAP-Y49F for the ULK1 LIR peptide were not different from wild-type proteins. The dissociation constant for LC3A-F52Y was 3.8 ± 0.5 μM, compared with 4.4 ± 0.5 μM for LC3A. The dissociation constant for the GABARAP-Y49F mutant was 0.42 ± 0.03 μM, compared with 0.3 ± 0.05 μM for GABARAP. The sensograms for LC3A-F52Y and GABARAP-Y49F were consistent with the wild-type and arginine mutant data, with the same trend of fast on and off rates (Figures 5.15A and 5.15C, replicate data shown in Figure 5.16). Thus, the presence or absence of the hydroxyl group had little effect on the binding affinity of LC3A and GABARAP for the ULK1-LIR.

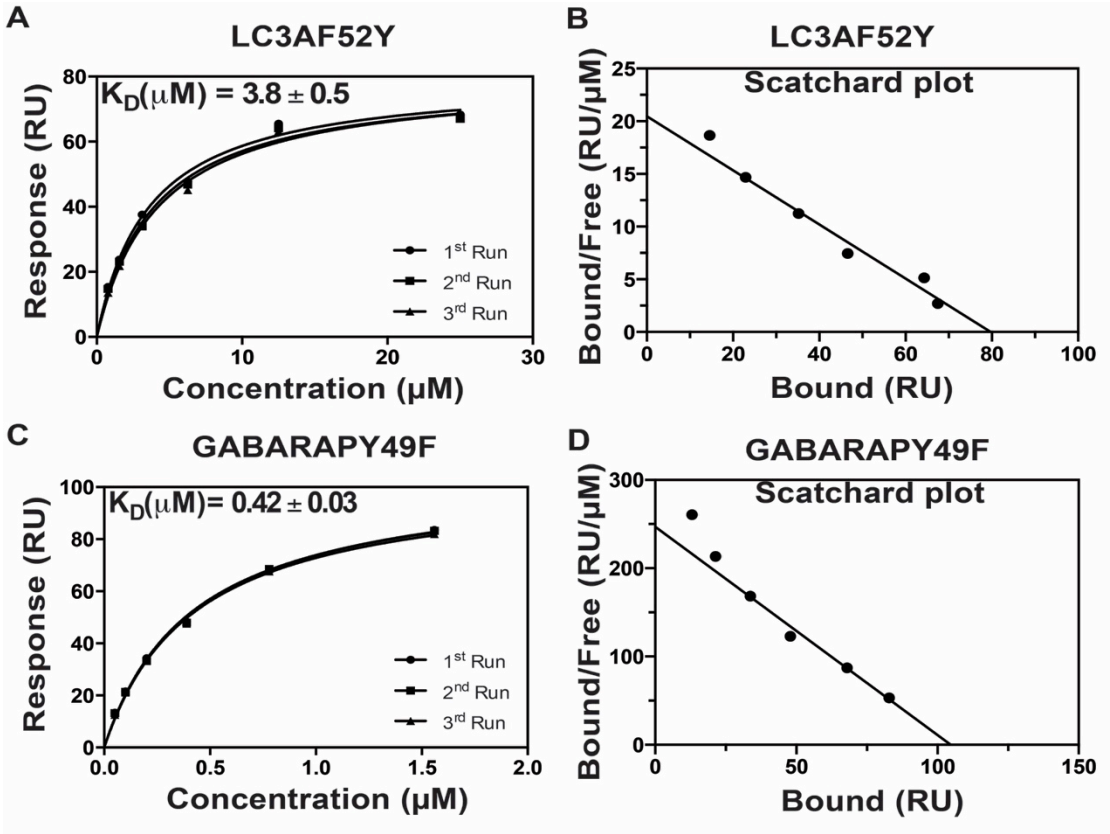


Figure 5.15: A & C. 1:1 Langmuir models for LC3-F52Y and GABARAP-Y49F fitted for triplicate data sets. The equilibrium binding constants were calculated from the response vs. concentration curve. There were no differences in K_D between LC3A-F52Y and LC3A-wt or between GABARAP-Y49F and GABARAP-wt. B & D. Scatchard plots for the average from the triplicate data show GABARAP-Y49F has higher binding affinity for the ULK1-LIR than LC3A-F52Y, which agrees with the 1:1 Langmuir model.

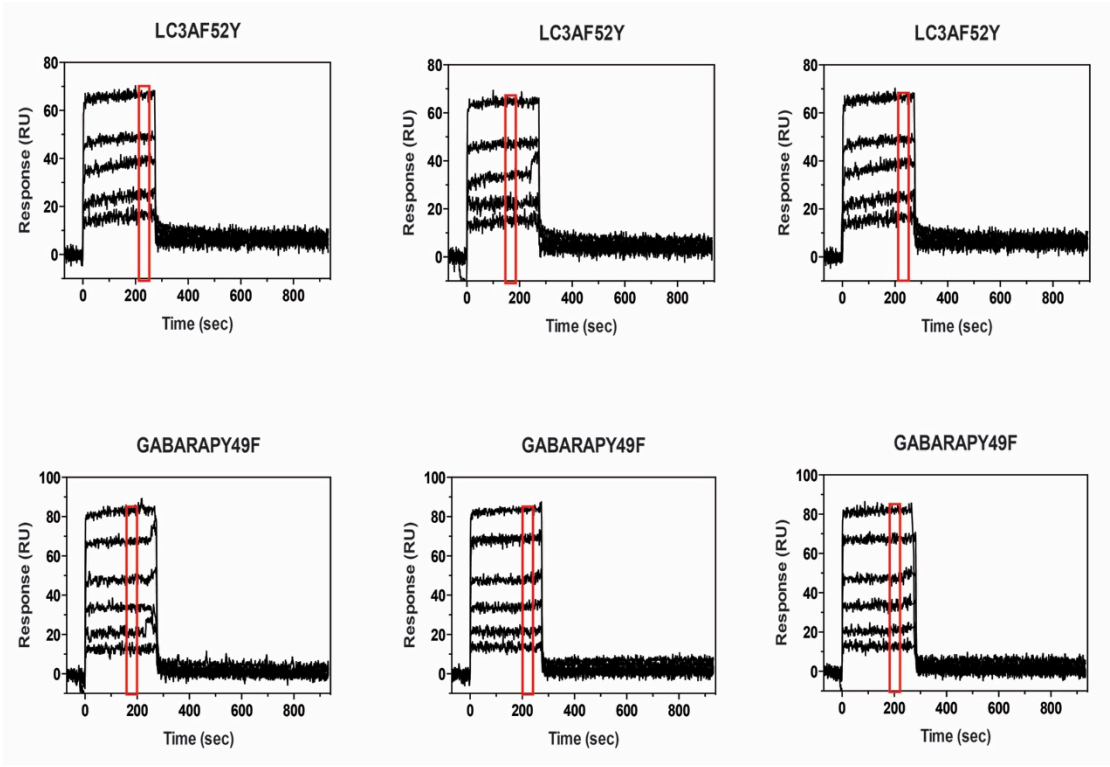


Figure 5.16: SPR response data collected in triplicate for LC3A-F52Y and GABARAP-Y49F. The response data is shown in black, and six concentrations of both LC3A-F52Y (25 μM, 12.5 μM, 6.25 μM, 3.125 μM, 1.56 μM and 0.78 μM) and GABARAP-Y49F (1.56 μM, 0.78 μM, 0.39 μM, 0.195 μM, 0.097 μM and 0.048 μM) were injected. The data points from the regions marked by red boxes were used in the 1:1 Langmuir model to calculate the binding affinity.

5.3 Summary

The GABARAP protein has a 14-fold higher binding affinity towards the ULK1 LIR than LC3A, as demonstrated by SPR. Although the binding affinities differ significantly, LC3A and GABARAP displayed a similar response profile (box shape), which matched the response profile reported by Suzuki *et al.* (2014). The difference in binding affinity, despite the close sequence similarity and similar binding response profile, suggested that the ULK1 LIR interaction mechanisms of GABARAP and LC3A are different.

The interaction of the ULK1 LIR motif with GABARAP was verified by solving the crystal structure of the ULK1 LIR fused to GABARAP. Structural

comparison of liganded and unliganded forms of GABARAP and LC3A demonstrated that LC3A has a more dynamic “W-site” hydrophobic pocket, which resembles an open-and-close mechanism, whereas the GABARAP “W-site” hydrophobic pocket remained in an open state regardless of ligand binding, suggesting a more stable and less dynamic “W-site”.

Molecular dynamic simulation suggested that LC3A was more dynamic than GABARAP. Interestingly, thermostability data supported the simulation data by demonstrating that GABARAP has a higher melting temperature than the more flexible LC3A. Further analysis of the “W-site” hydrophobic pocket suggested that the aromatic side chain of Phe357 in the ULK1 LIR made a more stable interaction with GABARAP than with LC3A. Taken together, the structural, SPR and MD simulation data suggest that the “W-site” hydrophobic pocket regulates the ULK1 LIR interaction differently in LC3A than in GABARAP. Further investigation of the role of Arg70 in LC3A and Arg67 in GABARAP suggested that the former plays a key role in the interaction with Met359 in the ULK1 LIR, whereas the latter does not have a role in LIR interaction.

Mutating the corresponding arginines in both LC3A (Arg70) and GABARAP (Arg67) to leucine affected the LIR binding affinities of the respective proteins differently. The R70L mutation in LC3A caused a significant 2.5-fold decrease in binding affinity compared with wild-type LC3A, whereas the R67L mutant of GABARAP displayed no difference in binding affinity compared with wild-type GABARAP. The presence or absence of a hydroxyl group in Tyr49 (GABARAP) or Phe52 (LC3A) did not have an effect on the binding affinity. The above observations suggested that arginines at corresponding positions in LC3A and GABARAP interact

differently with the ULK1 LIR, and that the charge of the side chain of residue 70, located at the hydrophobic pocket, is important for LC3A but not for GABARAP.

In conclusion, this chapter provides an explanation for the differences in binding affinities of the ULK1 LIR towards LC3A and GABARAP, despite the structural and sequence similarities between these two proteins. The overall architecture of the GABARAP structure allows a stable hydrophobic interaction of the aromatic amino acid residue from the LIR motif at the "W-site", whereas the dynamic nature of the LC3A "W-site" hydrophobic surface results in a less stable interaction with the aromatic amino acid from the LIR. Introducing amino acids to stabilise the LC3A "W-site" pocket and analysing their effects on binding affinity would provide more information on the role of the "W-site" hydrophobic pocket in preferential binding of the LIR motifs. To balance the weak hydrophobic LIR interaction in LC3A, Arg70 plays a key role by interacting with the LIR, as shown by the change in binding affinity when it is replaced by leucine. In contrast, the arginine at the corresponding position in GABARAP does not have any effect on the binding affinity. These findings show a clear difference in the interaction mechanisms of LC3A and GABARAP with the ULK1 LIR motif.

Chapter 6. CONCLUSIONS

In eukaryotic cells, degradation of unstable proteins, protein aggregates, and organelles through a complex cellular pathway called autophagy is important for maintaining homeostasis (Mizushima *et al.*, 2008). More than 32 autophagy-related proteins orchestrate the autophagy pathway through a complex protein-protein interaction network (Behrends *et al.*, 2010). Although considerable effort has gone into understanding this pathway, the interaction between the proteins involved in autophagy is not well understood.

The work in this thesis focused on two interactions: (i) the LIR-independent interaction between ATG13 and FIP200, which are components of the human autophagy initiation complex known as the ULK1 complex; and (ii) the LIR-dependent interaction between members of the human ATG8 protein family and the PLEKHM1 LIR peptide, and the LIR-dependent interaction between LC3A and GABARAP with the ULK1 LIR peptide.

6.1 Mapping the ATG13 binding region in FIP200.

FIP200 and ATG13 are components of the ULK1 initiation complex. However, it is not known how the two proteins interact with one another. Previous studies suggest that the ULK1-ATG13-FIP200 complex is important for the regulation of starvation-induced autophagy (Jung *et al.*, 2009, Chang & Neufeld, 2009). The working model is that, under nutrient-rich growth conditions, the serine/threonine kinase activity of ULK1 is supressed by the phosphorylation of ULK1 and ATG13 by mTOR. During starvation, inhibition of mTOR leads to phosphorylation of ATG13 and FIP200 by ULK1, which in turn leads to the induction

of autophagy (Chan, 2009). However, a knock-out study by Alers *et al.* (2011) that investigated the role of the ULK1-ATG13-FIP200 complex showed that the combined action of ATG13 and FIP200 was critical for the induction of autophagy, whereas the ULK1/ULK2 has no significant function in starvation-induced autophagy. Furthermore, in a co-immunoprecipitation assay, the amount of FIP200 that bound to ATG13 was higher than that of ULK1 and ULK2, suggesting that ATG13 binds FIP200 more tightly than ULK1/ULK2 (Jung *et al.*, 2009). Therefore, understanding the interaction between ATG13 and FIP200 is an important step towards a clear understanding of the function of the ULK1 initiation complex.

Previous work by Jung *et al.* (2009) showed that residues 384–517 in the C-terminal region of ATG13 bind FIP200, ULK1 and ULK2. Although the ATG13 binding sites on ULK1 (829–1051) and ULK2 (651–1036) have been mapped, the site on FIP200 that binds ATG13 is unknown. In this thesis, attempts were made to map the region of FIP200 that binds to ATG13. To narrow down the location of the binding region in FIP200, the full length FIP200 gene was divided into six parts, P1 (2–266), P2 (267–531), P3 (532–796), P4 (797–1061), P5 (1062–1328) and P6 (1327–1591), and each was cloned into the pCold™ TF expression vector to aid expression and solubility.

Pull-down assays were performed to map the ATG13 interaction site in FIP200. Because of the false positive interaction of FIP200 polypeptides with the GST tag, the pull-down assay analysis did not provide results that could be used for further analysis. A possible cause for this false positive could be that the FIP200 polypeptides were improperly folded, causing them to interact with the GST tag. Use of a different tag, such as maltose-binding protein, could be useful for improved purification of ATG13 and clearer results in the pull-down assay. Future experiments

to map the ATG13-binding region of FIP200 should include careful consideration of the regions to be truncated, further optimisation of the expression and purification parameters to ensure the quality of protein used in the pull-down assay is acceptable, verification of the folding of the FIP200 polypeptides and ATG13, and exploration of different techniques, such as SPR, to map the interaction.

6.2 GABARAP family proteins have higher binding affinity towards the PLEKHM1 LIR than LC3 family proteins.

GABARAP family proteins have higher binding affinity towards the PLEKHM1 LIR than LC3 family proteins. In Chapter 4, the interaction between the PLEKHM1 LIR and human ATG8 protein family members was investigated by solving the crystal structures of the PLEKHM1 LIR in complex with LC3A, LC3C, GABARAP and GABARAPL1. The crystal structure of the PLEKHM1 LIR in complex with LC3B (PDB code: 3X0W, (McEwan *et al.*, 2015)) was included in the analysis.

Recently, PLEKHM1 was identified as a human ATG8 interaction partner that facilitates fusion between autophagosomes and lysosomes (McEwan *et al.*, 2015). Interestingly, unpublished ITC data (refer to Chapter 4, Table 4.1) kindly provided by Vladimir Rogov from Prof. Ivan Dikic's lab shows that GABARAP family proteins have significantly higher binding affinity towards the PLEKHM1 LIR peptide (GABARAP: K_D = 0.55 μM and GABARPL1: K_D = 0.77 μM) than LC3 family proteins (LC3A: K_D = 4.22 μM, LC3B K_D = 6.33 μM and LC3C K_D = 3.45 μM). Similar findings were reported for the interaction between the human ATG8 family proteins with the members of the ULK1 initiation complex (ULK1, FIP200 and

ATG13), where GABARAP family proteins were shown to have higher binding affinity than LC3 family proteins (Alemu *et al.*, 2012).

Based on this information, it was predicted that the structural comparison of the PLEKHM1 LIR in complex with LC3A, LC3B, LC3C, GABARAP and GABARAPL1 would show significant differences in the interaction of the PLEKHM1 LIR across the five human ATG8 proteins, and provide a possible explanation for the difference in binding affinity between GABARAP proteins and LC3 proteins. However, all five structures of the PLEKHM1 LIR in complex with human ATG8 proteins were comparable, and interacted with the PLEKHM1 LIR in a similar manner (refer to Section 4.2.2). Despite this similarity across all five human ATG8 structures (LC3A, LC3B, LC3C, GABARAP and GABARAPL1), subtle differences may contribute to the tighter interaction of the PLEKHM1 LIR with the GABARAP proteins than with the LC3 proteins.

In particular, the interaction of the PLEKHM1 LIR residues Asn637 and Trp635 with the human ATG8 proteins was explored (refer to Section 4.2.3). There are several reasons why these electrostatic interactions show promise as an explanation for the higher binding affinity of the PLEKHM1 LIR towards GABARAP proteins (refer to Table 4.3 for summary of interactions and binding affinity). First, the presence of Arg28 in GABARAP and GABARAPL1, compared with lysine at the corresponding position in LC3 proteins, allows the side chain of Arg28 to structurally rearrange by 4 Å, and interact with Asn637 of the PLEKHM1 LIR (refer Figure 4.7). A similar interaction was found between Lys36 (without side chain structural re-arrangement) of LC3C, which has the highest binding affinity among LC3 proteins, and Asn637 of the PLEKHM1 LIR; however, this residue not found in other LC3 proteins. Second, the PLEKHM1 LIR residue Trp635 interacts with Glu17 in

GABARAP and GABARAPL1, whereas a similar interaction was not observed in LC3 proteins (refer Figure 4.7).

A survey of the Protein Data Bank for GABARAP family proteins in complex with other LIR motifs was conducted. In the case of the autophagy-linked FYVE protein (ALFY) LIR-GABARAP complex structure (a comparison of PLEKHM1 LIR-GABARAP and ALFY LIR-GABARAP complexes is shown in Figure 6.1A and B), the electrostatic interaction between GABARAP residues Lys24/Tyr25/Asp54 and ALFY residues Asp3344 and Tyr3351 may contribute to the binding affinity (Lystad *et al.*, 2014). However, residues at the corresponding positions in LC3A and LC3B (Gln26/His27/His57) did not interact with ALFY residues Asp3344 and Tyr3351 (Lystad *et al.*, 2014). The interaction of Asp3344 and Tyr3351 of the ALFY LIR with GABARAP, but not with LC3A and LC3B, is only a subtle difference but is proposed to have an effect on binding affinity with the LIR (Lystad *et al.*, 2014). Similarly, the interaction of Trp635 and Asn637 in the PLEKHM1 LIR with GABARAP proteins, which is either weak or absent in LC3 proteins (see Section 4.2.3), may contribute to the tighter binding of the PLEKHM1 LIR with GABARAP proteins than with LC3 proteins.

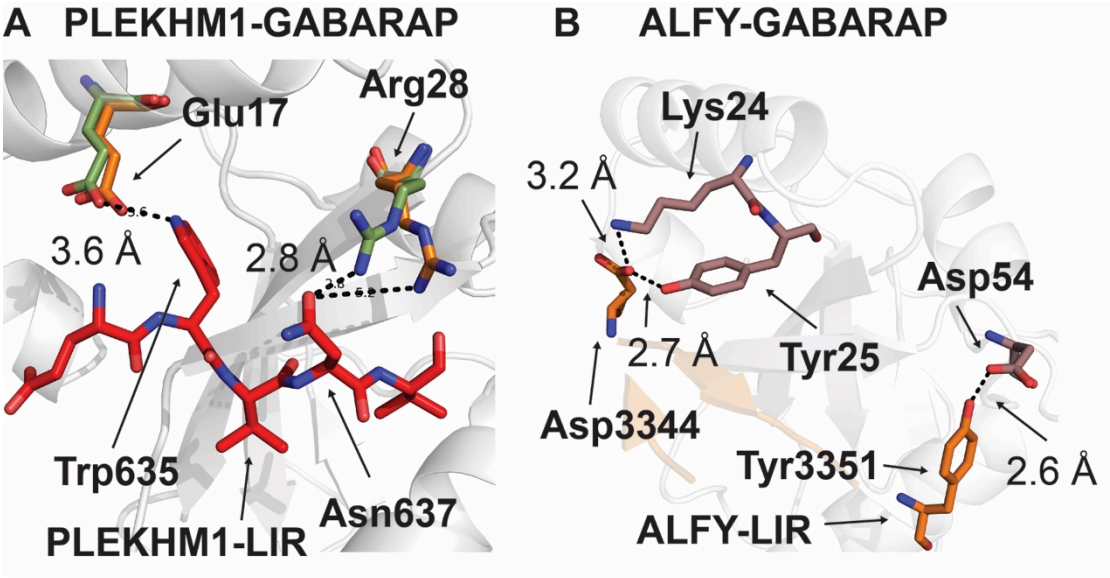


Figure 6.1: A. Interaction of the PLEKHM1 LIR with GABARAP. Comparison between liganded (green) and unliganded (orange) Glu17 shows that Trp635 of the PLEKHM1 LIR interacts with Glu17 in the liganded form. Similarly, Arg28 interacts with Asn637 of the PLEKHM1 LIR. B. Interaction of the ALFY LIR (orange) with GABARAP. Lys24 and Tyr25 of GABARAP interact with Asp3344 of the ALFY LIR, and Asp54 of GABARAP interacts with Tyr3351 of the ALFY LIR, demonstrating electrostatic interactions besides the hydrophobic interaction.

In addition to the differences in electrostatic interactions between LC3 and GABARAP proteins, there is a difference in the "W-site" hydrophobic pocket (refer to Section 4.2.4). In GABARAP proteins, the "W-site" hydrophobic pocket is deeper than in LC3 proteins, and has a conformation that can better accommodate the bulkier aromatic amino acid from the LIR motif (refer to Figure 4.8). This notable difference in the architecture of the "W-site" hydrophobic pocket between the two protein families may also contribute to the variation in binding affinity for the PLEKHM1 LIR across the members of the human ATG8 family of proteins.

Taken together, the above discussion provides a possible rationale for the higher binding affinity towards the PLEKHM1 LIR demonstrated by GABARAP family proteins. These interactions could be further probed by verifying the binding affinity of Glu17/Arg28 double mutants of GABARAP proteins towards the PLEKHM1 LIR.

The difference in binding affinity might also have implications for biological function. Because PLEKHM1 mediates the fusion of the lysosome to the autophagosome (called the autolysosome) (McEwan *et al.*, 2015), which occurs during the final stage of autophagosome formation, it is possible that PLEKHM1 has a preference for GABARAP proteins over LC3 proteins. This may be explained by the observations that GABARAP plays a role in the closure of the autophagosome double membrane vesicle, whereas LC3 plays a role during the initial stage of autophagosome formation (Weidberg *et al.*, 2010).

6.3 The "W-site" hydrophobic pocket and Arg70 in LC3A affect the interaction with the ULK1 LIR.

Chapter 5 describes the role of the "W-site" hydrophobic pocket during the interaction with the ULK1 LIR peptide and the involvement of Arg70 in LC3A and Arg67 in GABARAP. Behrends *et al.* (2010) demonstrated that 18 proteins with an LIR motif bound with lower affinity to LC3B-Arg70Ala than to the wild-type protein. In contrast, Arg67Ala in GABARAP (the equivalent mutation) had no effect on binding affinity. Therefore, in Chapter 5, amino acids Arg70 in LC3A and Arg67 in GABARAP were probed structurally and kinetically using SPR, MD simulation and X-ray crystallography, to understand the arginine-mediated mechanistic difference in LIR interaction between GABARAP and LC3A.

SPR experiments demonstrated that the GABARAP protein has 14-fold higher binding affinity towards the ULK1 LIR compared with LC3A (refer to Section 5.2.2). The difference in binding affinity, despite the close sequence similarity and similar binding response profile, suggests that the GABARAP and LC3A interact differently with the ULK1 LIR. Structural comparison of the liganded (ULK1-GABARAP and

previously solved ATG13-LC3A) and unliganded states of GABARAP and LC3A showed both open and closed conformations of the "W-site" hydrophobic pocket in the LC3A structure, whereas the "W-site" of GABARAP had an open conformation, regardless of its ligand-binding state (refer to Section 5.2.3 and Figure 5.7).

Molecular dynamics simulations suggested that LC3A was more dynamic than GABARAP, and that ULK1 LIR residue Phe357 interacting with the "W-site" hydrophobic pocket is more dynamic in LC3A than in GABARAP, suggesting that GABARAP makes a more stable interaction with the ULK1 LIR. The simulation also suggested the interaction of Arg70 in LC3A with Met359 of the ULK1 LIR. This provided key information that the side chain charge of Arg70 could be important for the LIR interaction in LC3A.

It was therefore predicted that when Arg70 in LC3A and Arg67 in GABARAP were mutated to leucine, it would affect the binding affinity towards the ULK1 LIR differently for the two proteins. It was further hypothesised that the mutation would have opposite effects on the binding affinity of LC3A and GABARAP with the ULK1-LIR, and this effect on binding affinity would be caused by the loss of the charge in the side chain of the arginine. Using SPR, the effect of mutating Arg70 in LC3A and Arg67 in GABARAP to leucine was probed (refer to Section 5.2.6.2). LC3A-R70L displayed a significant 2.5-fold decrease in binding affinity compared with the wild-type LC3A, whereas the binding affinity of GABARAP-R67L remained unchanged from wild-type GABARAP (refer Figure 5.13). This suggests that Arg70 in LC3A plays a role in LIR interaction, but Arg67 in GABARAP does not.

The key finding for this section was that the "W-site" hydrophobic pocket plays an important role in the preferential binding of the LIR motif. This is

demonstrated by: 1) SPR data shows higher binding affinity for the ULK1 LIR towards GABARAP than to LC3A, 2) structural comparison of ULK1 LIR-GABARAP and ATG13 LIR-LC3A shows a more dynamic LC3A "W-site" hydrophobic pocket between its LIR-bound and -unbound states, 3) molecular dynamics simulation shows that interaction of F357 of the ULK1 LIR at the "W-site" in LC3A and GABARAP is more dynamic in the former than the in the latter, suggesting a less stable interaction in LC3A.

A second finding is that there may be differences in the mechanisms by which LC3A and GABARAP bind to the LIR peptide. This is shown by: 1) the observation that mutation of Arg70 in LC3A affects the binding affinity for the ULK1 LIR, but mutation at the equivalent residue in GABARAP (Arg67) does not, and 2) the MD simulation suggests an electrostatic interaction between Arg70 of LC3A and Met359 of the ULK1 LIR, whereas no such interaction is observed between GABARAP Arg67 and Met359 of the ULK1 LIR.

In future, performing SPR to measure binding affinities of LC3A with Phe52 mutated to a non-aromatic amino acid or Arg70/Phe52 double mutants will shed some light on the importance of Phe52 in LC3A. Performing cell culture studies on GABARAP-R67L and LC3A-R70L mutants would provide evidence on the biological functions of Arg70 in LC3A and Arg67 in GABARAP. Solving the crystal structure of GABARAP-R67L and LC3A-R70L in complex with the ULK1 LIR and comparing them ULK1 LIR-bound wild-type structures will also provide insight into the structural changes occurring in Arg70 during LIR interaction.

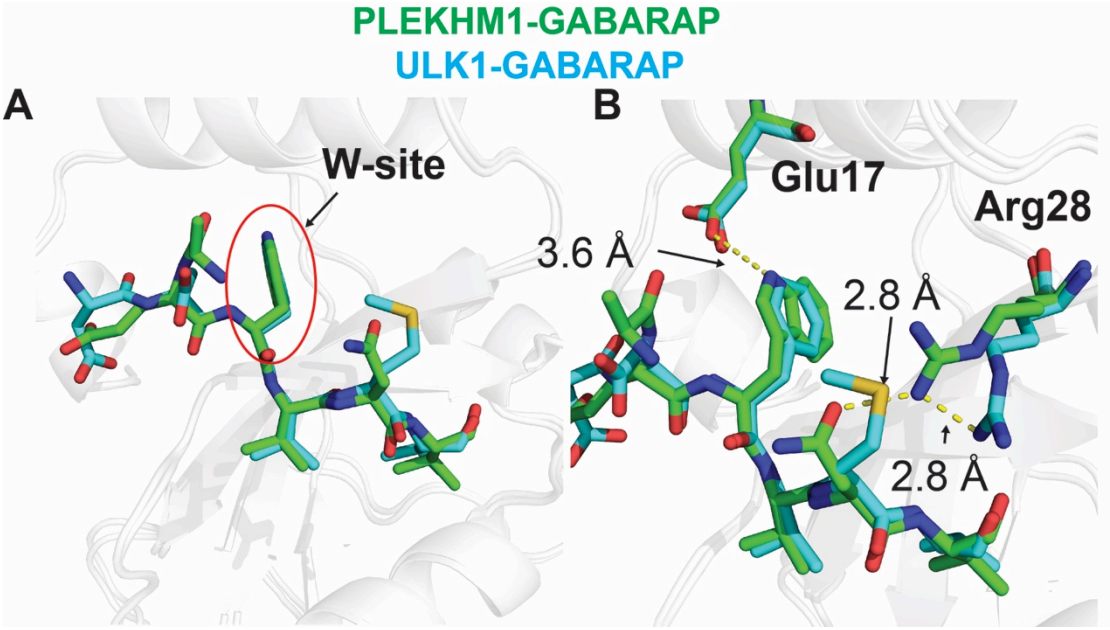


Figure 6.2: A. Interaction of the PLEKHM1 LIR (green) and the ULK1 LIR (cyan) with GABARAP. Both LIRs interact in similar binding conformations, especially Trp635 of PLEKHM1 and Phe357 of ULK1. B. Glu17 and Arg28 of GABARAP interact with Trp635 and Asn637 of the PLEKHM1 LIR, whereas neither of these interactions are observed in the ULK1 LIR-GABARAP structure.

Overall, two types of LIR motifs were probed in this study: the PLEKHM1 LIR, from a selective autophagy receptor, and the ULK1 LIR, from the macroautophagy pathway. Studies showed the importance of the "W-site" hydrophobic pocket in preferential interaction with both LIR motifs. This was demonstrated by the interaction of Glu17 at the "W-site" of GABARAP with Trp635 of the PLEKHM1 LIR, and a the less dynamic interaction between F357 of the ULK1 LIR and the GABARAP "W-site". Surprisingly, the binding conformation of Trp635 in the PLEKHM1 LIR and Phe357 of the ULK1 LIR is very similar, suggesting that the interaction at the "W-site" is important in GABARAP (Figure 6.2A). However, the interaction between GABARAP Arg28 and the PLEKHM1 LIR Asn637 was not found in the ULK1 LIR-GABARAP structure (Figure 6.2B). The work in this thesis provides initial insights into differences in LIR interaction mechanisms for LC3 and GABARAP proteins, which can be used in future studies to understand the need for the six ATG8 homologues in humans compared with only one ATG8 in yeast.

Chapter 7. REFERENCES

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *PHENIX: A comprehensive Python-based system for macromolecular structure solution*, *Acta Crystallographica Section D: Biological Crystallography* **66**, 213-221.

Aguado, C., Sarkar, S., Korolchuk, V. I., Criado, O., Vernia, S., Boya, P., Sanz, P., de Córdoba, S. R., Knecht, E. & Rubinsztein, D. C. (2010). *Laforin, the most common protein mutated in Lafora disease, regulates autophagy*, *Human Molecular Genetics* **19**, 2867-2876.

Ahlberg, J. & Glaumann, H. (1985). *Uptake - microautophagy - and degradation of exogenous proteins by isolated rat liver lysosomes. Effects of pH, ATP, and inhibitors of proteolysis*, *Experimental and Molecular Pathology* **42**, 78-88.

Alemu, E. A., Lamark, T., Torgersen, K. M., Birgisdottir, A. B., Larsen, K. B., Jain, A., Olsvik, H., Øvervatn, A., Kirkin, V. & Johansen, T. (2012). *ATG8 family proteins act as scaffolds for assembly of the ULK complex: Sequence requirements for LC3-interacting region (LIR) motifs*, *Journal of Biological Chemistry* **287**, 39275-39290.

Alers, S., Löffler, A. S., Paasch, F., Dieterle, A. M., Keppeler, H., Lauber, K., Campbell, D. G., Fehrenbacher, B., Schaller, M., Wesselborg, S. & Stork, B. (2011). *Atg13 and FIP200 act independently of Ulk1 and Ulk2 in autophagy induction*, *Autophagy* **7**, 1424-1433.

Baneyx, F. (1999). *Recombinant protein expression in Escherichia coli*, *Current Opinion in Biotechnology* **10**, 411-421.

Baumeister, W., Walz, J., Zühl, F. & Seemüller, E. (1998). *The proteasome: Paradigm of a self-compartmentalizing protease*, *Cell* **92**, 367-380.

Bavro, V. N., Sola, M., Bracher, A., Kneussel, M., Betz, H. & Weissenhorn, W. (2002). *Crystal structure of the GABAA-receptor-associated protein, GABARAP*, *EMBO reports* **3**, 183-189.

Behrends, C., Sowa, M. E., Gygi, S. P. & Harper, J. W. (2010). *Network organization of the human autophagy system*, *Nature* **466**, 68-76.

Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. (1995). *GROMACS: A message-passing parallel molecular dynamics implementation*, *Computer Physics Communications* **91**, 43-56.

Berger, Z., Ravikumar, B., Menzies, F. M., Oroz, L. G., Underwood, B. R., Pangalos, M. N., Schmitt, I., Wullner, U., Evert, B. O., O'Kane, C. J. & Rubinsztein, D. C. (2006). *Rapamycin alleviates toxicity of different aggregate-prone proteins*, *Human Molecular Genetics* **15**, 433-442.

Bjørkøy, G., Lamark, T., Brech, A., Outzen, H., Perander, M., Øvervatn, A., Stenmark, H. & Johansen, T. (2005). *p62/SQSTM1 forms protein aggregates degraded by autophagy and has a protective effect on huntingtin-induced cell death*, *The Journal of Cell Biology* **171**, 603-614.

Bursch, W., Ellinger, A., Kienzl, H., Török, L., Pandey, S., Sikorska, M., Walker, R. & Hermann, R. S. (1996). *Active cell death induced by the anti-estrogens tamoxifen and ICI 164 384 in human mammary carcinoma cells (MCF-7) in culture: The role of autophagy*, *Carcinogenesis* **17**, 1595-1607.

Campbell, G. R. & Spector, S. A. (2011). *Hormonally active vitamin D3 (1α,25-dihydroxycholecalciferol) triggers autophagy in human macrophages that inhibits HIV-1 infection*, *Journal of Biological Chemistry* **286**, 18890-18902.

Caramés, B., Taniguchi, N., Otsuki, S., Blanco, F. J. & Lotz, M. (2010). *Autophagy is a protective mechanism in normal cartilage, and its aging-related loss is linked with cell death and osteoarthritis*, *Arthritis and Rheumatism* **62**, 791-801.

Chan, E. Y. (2009). *mTORC1 Phosphorylates the ULK1-mAtg13-FIP200 Autophagy Regulatory Complex*, *Science Signaling* **2**, pe51-pe51.

Chan, E. Y. W., Kir, S. & Tooze, S. A. (2007). *siRNA Screening of the Kinome Identifies ULK1 as a Multidomain Modulator of Autophagy*, *Journal of Biological Chemistry* **282**, 25464-25474.

Chang, Y.-Y. & Neufeld, T. P. (2009). *An Atg1/Atg13 Complex with Multiple Roles in TOR-mediated Autophagy Regulation*, *Molecular Biology of the Cell* **20**, 2004-2014.

Chano, T., Ikegawa, S., Kontani, K., Okabe, H., Baldini, N. & Saeki, Y. (2002). *Identification of RB1CC1, a novel human gene that can induce RB1 in various human cells*, *Oncogene* **21**, 1295-1298.

Codogno, P. & Meijer, A. J. (2010). *Autophagy: A potential link between obesity and insulin resistance*, *Cell Metabolism* **11**, 449-451.

Cooper, G. M. (2000a). *Protein Degradation*. Sunderland (MA): Sinauer Associates.

Cooper, G. M. (2000b). *Lysosomes*. Sunderland (MA): Sinauer Associates.

Cuervo, A. M. (2010). *Chaperone-mediated autophagy: Selectivity pays off*, *Trends in Endocrinology and Metabolism* **21**, 142-150.

Cuervo, A. M. & Dice, J. F. (1996). *A receptor for the selective uptake and degradation of proteins by lysosomes*, *Science* **273**, 501-503.

Cuervo, A. M. & Wong, E. (2014). *Chaperone-mediated autophagy: Roles in disease and aging*, *Cell Research* **24**, 92-104.

De Duve, C. (1963). *The lysosome*, *Scientific American* **208**, 64-72.

De Duve, C. & Wattiaux, R. (1966). *Functions of lysosomes*, *Annual Review of Physiology* **28**, 435-492.

De Kreutzenberg, S. V., Ceolotto, G., Papparella, I., Bortoluzzi, A., Semplicini, A., Dalla Man, C., Cobelli, C., Fadini, G. P. & Avogaro, A. (2010). *Downregulation of the longevity-associated protein sirtuin 1 in insulin resistance and metabolic syndrome: Potential biochemical mechanisms*, *Diabetes* **59**, 1006-1015.

DeSelm, C. J., Miller, B. C., Zou, W., Beatty, W. L., van Meel, H., Takahata, Y., Klumperman, J., Tooze, S. A., Teitelbaum, S. L. & Virgin, H. W. (2011). *Autophagy Proteins Regulate the Secretory Component of Osteoclastic Bone Resorption*, *Dev. Cell* **21**, 966-974.

Dos, D. S., Ali, S. M., Kim, D.-H., Guertin, D. A., Latek, R. R., Erdjument-Bromage, H., Tempst, P. & Sabatini, D. M. (2004). *Rictor, a Novel Binding Partner of mTOR, Defines a Rapamycin-Insensitive and Raptor-Independent Pathway that Regulates the Cytoskeleton*, *Current Biology* **14**, 1296-1302.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Features and development of Coot*, *Acta Crystallographica Section D: Biological Crystallography* **66**, 486-501.

English, L., Chemali, M., Duron, J., Rondeau, C., Laplante, A., Gingras, D., Alexander, D., Leib, D., Norbury, C., Lippé, R. & Desjardins, M. (2009). *Autophagy enhances the presentation of endogenous viral antigens on MHC class I molecules during HSV-1 infection*, *Nature Immunology* **10**, 480-487.

Evans, P. (2006). *Scaling and assessment of data quality*, *Acta Crystallographica Section D* **62**, 72-82.

Evans, P. R. & Murshudov, G. N. (2013). *How good are my data and what is the resolution?*, *Acta Crystallographica Section D: Biological Crystallography* **69**, 1204-1214.

Fischer, M. J. E. (2010). *Amine coupling through EDC/NHS: a practical approach*, *Methods Mol. Biol. (Totowa, NJ, U. S.)* **627**, 55-73.

Fleming, A., Noda, T., Yoshimori, T. & Rubinsztein, D. C. (2011). *Chemical modulators of autophagy as biological probes and potential therapeutics*, *Nature Chemical Biology* **7**, 9-17.

Gan, B., Melkoumian, Z. K., Wu, X., Guan, K. L. & Guan, J. L. (2005). *Identification of FIP200 interaction with the TSC1-TSC2 complex and its role in regulation of cell size control*, *Journal of Cell Biology* **170**, 379-389.

Gan, B., Peng, X., Nagy, T., Alcaraz, A., Gu, H. & Guan, J. L. (2006). *Role of FIP200 in cardiac and liver development and its regulation of TNFα and TSC-mTOR signaling pathways*, *Journal of Cell Biology* **175**, 121-133.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M., Appel, R. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. Walker, pp. 571-607: Humana Press.

Genau, H., Huber, J., Baschieri, F., Akutsu, M., Dötsch, V., Farhan, H., Rogov, V. & Behrends, C. (2015). *CUL3-KBTBD6/KBTBD7Ubiquitin Ligase Cooperates with GABARAP Proteins to Spatially Restrict TIAM1-RAC1 Signaling*, *Molecular Cell* **57**, 995-1010.

Hahn, M. (1898). *Chemistry of vegetable physiology and agriculture*, *Journal of the Chemical Society, Abstracts* **74**, B245-B252.

Hailey, D. W., Rambold, A. S., Satpute-Krishnan, P., Mitra, K., Sougrat, R., Kim, P. K. & Lippincott-Schwartz, J. (2010). *Mitochondria Supply Membranes for Autophagosome Biogenesis during Starvation*, *Cell* **141**, 656-667.

Hara, K., Maruki, Y., Long, X., Yoshino, K.-i., Oshiro, N., Hidayat, S., Tokunaga, C., Avruch, J. & Yonezawa, K. (2002). *Raptor, a Binding Partner of Target of Rapamycin (TOR), Mediates TOR Action*, *Cell* **110**, 177-189.

Hara, T., Takamura, A., Kishi, C., Iemura, S. I., Natsume, T., Guan, J. L. & Mizushima, N. (2008). *FIP200, a ULK-interacting protein, is required for autophagosome formation in mammalian cells*, *Journal of Cell Biology* **181**, 497-510.

Harding, T. M., Morano, K. A., Scott, S. V. & Klionsky, D. J. (1995). *Isolation and characterization of yeast mutants in the cytoplasm to vacuole protein targeting pathway*, *Journal of Cell Biology* **131**, 591-602.

Hayashi-Nishino, M., Fujita, N., Noda, T., Yamaguchi, A., Yoshimori, T. & Yamamoto, A. (2009). *A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation*, *Nature Cell Biology* **11**, 1433-1437.

He, C., Bassik, M. C., Moresi, V., Sun, K., Wei, Y., Zou, Z., An, Z., Loh, J., Fisher, J., Sun, Q., Korsmeyer, S., Packer, M., May, H. I., Hill, J. A., Virgin, H. W., Gilpin, C., Xiao, G., Bassel-Duby, R., Scherer, P. E. & Levine, B. (2012). *Exercise-induced BCL2-regulated autophagy is required for muscle glucose homeostasis*, *Nature* **481**, 511-515.

He, C. & Levine, B. (2010). *The Beclin 1 interactome*, *Current Opinion in Cell Biology* **22**, 140-149.

He, H., Dang, Y., Dai, F., Guo, Z., Wu, J., She, X., Pei, Y., Chen, Y., Ling, W., Wu, C., Zhao, S., Liu, J. O. & Yu, L. (2003). *Post-translational modifications of three members of the human MAP1LC3 family and detection of a novel type of modification for MAP1LC3B*, *Journal of Biological Chemistry* **278**, 29278-29287.

Heinemeyer, W., A.kleinschmidt, J., Saidowsky, J., Escher, C. & H.wolf, D. (1991). *Proteinase yscE, the yeast proteasome/multicatalytic-multifunctional proteinase: Mutants unravel its function in stress induced proteolysis and uncover its necessity for cell survival*, *EMBO Journal* **10**, 555-562.

Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. (2008). *GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation*, *Journal of Chemical Theory and Computation* **4**, 435-447.

Hochstrasser, M. (2009). *Origin and function of ubiquitin-like proteins*, *Nature* **458**, 422-429.

Hosokawa, N., Hara, T., Kaizuka, T., Kishi, C., Takamura, A., Miura, Y., Iemura, S. I., Natsume, T., Takehana, K., Yamada, N., Guan, J. L., Oshiro, N. & Mizushima, N. (2009). *Nutrient-dependent mTORCl association with the ULK1-Atg13-FIP200 complex required for autophagy*, *Molecular Biology of the Cell* **20**, 1981-1991.

Huang, J. & Manning, B. D. (2008). *The TSC1–TSC2 complex: a molecular switchboard controlling cell growth*, *The Biochemical journal* **412**, 179-190.

Humphrey, W., Dalke, A. & Schulten, K. (1996). *VMD: Visual molecular dynamics*, *Journal of Molecular Graphics* **14**, 33-38.

Hutchins, M. U. & Klionsky, D. J. (2001). *Vacuolar Localization of Oligomeric α-Mannosidase Requires the Cytoplasm to Vacuole Targeting and Autophagy Pathway Components in Saccharomyces cerevisiae*, *Journal of Biological Chemistry* **276**, 20491-20498.

Ichimura, Y., Kominami, E., Tanaka, K. & Komatsu, M. (2008). *Selective turnover of p62/A170/SQSTM1 by autophagy*, *Autophagy* **4**, 1063-1066.

Ichimura, Y., Kumanomidou, T., Sou, Y.-s., Mizushima, T., Ezaki, J., Ueno, T., Kominami, E., Yamane, T., Tanaka, K. & Komatsu, M. (2008). *Structural Basis for Sorting Mechanism of p62 in Selective Autophagy*, *J. Biol. Chem.* **283**, 22847-22857.

Igloi, G. L. (2001). *Cloning, expression patterns, and chromosome localization of three human and two mouse homologues of GABAA receptor-associated protein*, *Genomics* **74**, 408-413.

Janku, F., McConkey, D. J., Hong, D. S. & Kurzrock, R. (2011). *Autophagy as a target for anticancer therapy*, *Nature Reviews Clinical Oncology* **8**, 528-539.

Johansen, T. & Lamark, T. (2011). *Selective autophagy mediated by autophagic adapter proteins*, *Autophagy* **7**, 279-296.

Jung Ch Fau - Jun, C. B., Jun Cb Fau - Ro, S.-H., Ro Sh Fau - Kim, Y.-M., Kim Ym Fau - Otto, N. M., Otto Nm Fau - Cao, J., Cao J Fau - Kundu, M., Kundu M Fau - Kim, D.-H. & Kim, D. H. (2009). *ULK-Atg13-FIP200 complexes mediate mTOR signaling to the autophagy machinery*, *Molecular Biology of the Cell* **20**.

Jung, C. H., Jun, C. B., Ro, S.-H., Kim, Y.-M., Otto, N. M., Cao, J., Kundu, M. & Kim, D.-H. (2009). *ULK-Atg13-FIP200 Complexes Mediate mTOR Signaling to the Autophagy Machinery*, *Molecular Biology of the Cell* **20**, 1992-2003.

Kabeya, Y., Kamada, Y., Baba, M., Takikawa, H., Sasaki, M. & Ohsumi, Y. (2005). *Atg17 functions in cooperation with Atg1 and Atg13 in yeast autophagy*, *Molecular Biology of the Cell* **16**, 2544-2553.

Kabeya, Y., Mizushima, N., Ueno, T., Yamamoto, A., Kirisako, T., Noda, T., Kominami, E., Ohsumi, Y. & Yoshimori, T. (2000). *LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing*, *EMBO Journal* **19**, 5720-5728.

Kabeya, Y., Mizushima, N., Yamamoto, A., Oshitani-Okamoto, S., Ohsumi, Y. & Yoshimori, T. (2004). *LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation*, *Journal of Cell Science* **117**, 2805-2812.

Kabsch, W. (2010a). *XDS*, *Acta Crystallographica Section D: Biological Crystallography* **66**, 125-132.

Kabsch, W. (2010b). *Integration, scaling, space-group assignment and post-refinement*, *Acta Crystallographica Section D: Biological Crystallography* **66**, 133-144.

Kamada, Y., Funakoshi, T., Shintani, T., Nagano, K., Ohsumi, M. & Ohsumi, Y. (2000). *Tor-mediated induction of autophagy via an Apg1 protein kinase complex*, *Journal of Cell Biology* **150**, 1507-1513.

Kaushik, S. & Cuervo, A. M. (2012). *Chaperone-mediated autophagy: a unique way to enter the lysosome world*, *Trends in cell biology* **22**, 407-417.

Kim, J., Kundu, M., Viollet, B. & Guan, K. L. (2011). *AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1*, *Nature Cell Biology* **13**, 132-141.

Kirisako, T., Baba, M., Ishihara, N., Miyazawa, K., Ohsumi, M., Yoshimori, T., Noda, T. & Ohsumi, Y. (1999). *Formation Process of Autophagosome Is Traced with Apg8/Aut7p in Yeast*, *The Journal of Cell Biology* **147**, 435-446.

Kirkin, V., Lamark, T., Sou, Y. S., Bjørkøy, G., Nunn, J. L., Bruun, J. A., Shvets, E., McEwan, D. G., Clausen, T. H., Wild, P., Bilusic, I., Theurillat, J. P., Øvervatn, A., Ishii, T., Elazar, Z., Komatsu, M., Dikic, I. & Johansen, T. (2009). *A Role for NBR1 in Autophagosomal Degradation of Ubiquitinated Substrates*, *Molecular Cell* **33**, 505-516.

Klionsky, D. J., Cuervo, A. M., Dunn Jr, W. A., Levine, B., Van Der Klei, I. & Seglen, P. O. (2007). *How shall i eat thee?*, *Autophagy* **3**, 413-416.

Knight, D., Harris, R., McAlister, M. S. B., Phelan, J. P., Geddes, S., Moss, S. J., Driscoll, P. C. & Keep, N. H. (2002). *The X-ray Crystal Structure and Putative Ligand-derived Peptide Binding Properties of γ-Aminobutyric Acid Receptor Type A Receptor-associated Protein*, *Journal of Biological Chemistry* **277**, 5556-5561.

Kunzt, J. B., Schwarz, H. & Mayer, A. (2004). *Determination of Four Sequential Stages during Microautophagy in Vitro*, *Journal of Biological Chemistry* **279**, 9987-9996.

Lang, T., Schaeffeler, E., Bernreuther, D., Bredschneider, M., Wolf, D. H. & Thumm, M. (1998). *Aut2p and Aut7p, two novel microtubule-associated proteins are essential for delivery of autophagic vesicles to the vacuole*, *EMBO Journal* **17**, 3597-3607.

Laplante, M. & Sabatini, David M. (2012). *mTOR Signaling in Growth Control and Disease*, *Cell* **149**, 274-293.

Lecker, S. H., Goldberg, A. L. & Mitch, W. E. (2006). *Protein degradation by the ubiquitin-proteasome pathway in normal and disease states*, *Journal of the American Society of Nephrology* **17**, 1807-1819.

Lee, H. K., Mattei, L. M., Steinberg, B. E., Alberts, P., Lee, Y. H., Chervonsky, A., Mizushima, N., Grinstein, S. & Iwasaki, A. (2010). *In Vivo Requirement for Atg5 in Antigen Presentation by Dendritic Cells*, *Immunity* **32**, 227-239.

Lee, J. H., Yu, W. H., Kumar, A., Lee, S., Mohan, P. S., Peterhoff, C. M., Wolfe, D. M., Martinez-Vicente, M., Massey, A. C., Sovak, G., Uchiyama, Y., Westaway, D., Cuervo, A. M. & Nixon, R. A. (2010). *Lysosomal proteolysis and autophagy require presenilin 1 and are disrupted by Alzheimer-related PS1 mutations*, *Cell* **141**, 1146-1158.

Lee, J. S., Li, Q., Lee, J. Y., Lee, S. H., Jeong, J. H., Lee, H. R., Chang, H., Zhou, F. C., Gao, S. J., Liang, C. & Jung, J. U. (2009). *FLIP-mediated autophagy regulation in cell death control*, *Nature Cell Biology* **11**, 1355-1362.

Levine, B., Mizushima, N. & Virgin, H. W. (2011). *Autophagy in immunity and inflammation*, *Nature* **469**, 323-335.

Li, W. W., Li, J. & Bao, J. K. (2012). *Microautophagy: Lesser-known self-eating*, *Cellular and Molecular Life Sciences* **69**, 1125-1136.

Li, Y., Wang, L. X., Yang, G., Hao, F., Urba, W. J. & Hu, H. M. (2008). *Efficient cross-presentation depends on autophagy in tumor cells*, *Cancer Research* **68**, 6889-6895.

Liang, X. H., Jackson, S., Seaman, M., Brown, K., Kempkes, B., Hibshoosh, H. & Levine, B. (1999). *Induction of autophagy and inhibition of tumorigenesis by beclin 1*, *Nature* **402**, 672-676.

Liang, X. H., Kleeman, L. K., Jiang, H. H., Gordon, G., Goldman, J. E., Berry, G., Herman, B. & Levine, B. (1998). *Protection against fatal sindbis virus encephalitis by Beclin, a novel Bcl-2-interacting protein*, *Journal of Virology* **72**, 8586-8596.

Lindahl, E., Hess, B. & van der Spoel, D. (2001). *GROMACS 3.0: a package for molecular simulation and trajectory analysis*, *Molecular modeling annual* **7**, 306-317.

Lipinski, M. M., Zheng, B., Lu, T., Yan, Z., Py, B. F., Ng, A., Xavier, R. J., Li, C., Yankner, B. A., Scherzer, C. R. & Yuan, J. (2010). *Genome-wide analysis reveals mechanisms modulating autophagy in normal brain aging and in Alzheimer's disease*, *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14164-14169.

Liu, K. & Czaja, M. J. (2013). *Regulation of lipid stores and metabolism by lipophagy*, *Cell Death and Differentiation* **20**, 3-11.

Lum, J. J., Bauer, D. E., Kong, M., Harris, M. H., Li, C., Lindsten, T. & Thompson, C. B. (2005). *Growth factor regulation of autophagy and cell survival in the absence of apoptosis*, *Cell* **120**, 237-248.

Lystad, A. H., Ichimura, Y., Takagi, K., Yang, Y., Pankiv, S., Kanegae, Y., Kageyama, S., Suzuki, M., Saito, I., Mizushima, T., Komatsu, M. & Simonsen, A. (2014). *Structural determinants in GABARAP required for the selective binding and recruitment of ALFY to LC3B-positive structures*, *EMBO Rep.* **15**, 557-565.

Mann, S. S. & Hammarback, J. A. (1994). *Molecular characterization of light chain 3. A microtubule binding subunit of MAP1A and MAP1B*, *Journal of Biological Chemistry* **269**, 11492-11497.

Mariño, G., Salvador-Montoliu, N., Fueyo, A., Knecht, E., Mizushima, N. & López-Otín, C. (2007). *Tissue-specific autophagy alterations and increased tumorigenesis in mice deficient in Atg4C/autophagin-3*, *Journal of Biological Chemistry* **282**, 18573-18583.

Matsuura, A., Tsukada, M., Wada, Y. & Ohsumi, Y. (1997). *Apg1p, a novel protein kinase required for the autophagic process in Saccharomyces cerevisiae*, *Gene* **192**, 245-250.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *Phaser crystallographic software*, *Journal of Applied Crystallography* **40**, 658-674.

McEwan, David G., Popovic, D., Gubas, A., Terawaki, S., Suzuki, H., Stadel, D., Coxon, Fraser P., Miranda de Stegmann, D., Bhogaraju, S., Maddi, K., Kirchof, A., Gatti, E., Helfrich, Miep H., Wakatsuki, S., Behrends, C., Pierre, P. & Dikic, I. (2015). *PLEKHM1 Regulates Autophagosome-Lysosome Fusion through HOPS Complex and LC3/GABARAP Proteins*, *Molecular Cell* **57**, 39-54.

Melkoumian, Z. K., Peng, X., Gan, B., Wu, X. & Guan, J. L. (2005). *Mechanism of cell cycle regulation by FIP200 in human breast cancer cells*, *Cancer Research* **65**, 6676-6684.

Menzies, F. M., Huebener, J., Renna, M., Bonin, M., Riess, O. & Rubinsztein, D. C. (2010). *Autophagy induction reduces mutant ataxin-3 levels and toxicity in a mouse model of spinocerebellar ataxia type 3*, *Brain* **133**, 93-104.

Metcalf, D. J., García-Arencibia, M., Hochfeld, W. E. & Rubinsztein, D. C. (2012). *Autophagy and misfolded proteins in neurodegeneration*, *Experimental Neurology* **238**, 22-28.

Mitch, W. E. & Goldberg, A. L. (1996). *Mechanisms of disease: Mechanisms of muscle wasting: The role of the ubiquitin-proteasome pathway*, *New England Journal of Medicine* **335**, 1897-1905.

Mizushima, N., Levine, B., Cuervo, A. M. & Klionsky, D. J. (2008). *Autophagy fights disease through cellular self-digestion*, *Nature* **451**, 1069-1075.

Mizushima, N., Sugita, H., Yoshimori, T. & Ohsumi, Y. (1998). *A new protein conjugation system in human: The counterpart of the yeast Apg12p conjugation system essential for autophagy*, *Journal of Biological Chemistry* **273**, 33889-33892.

Mizushima, N., Yoshimori, T. & Ohsumi, Y. (2011). The role of atg proteins in autophagosome formation, Vol. 27, pp. 107-132.

Morselli, E., Galluzzi, L., Kepp, O., Mariño, G., Michaud, M., Vitale, I., Maiuri, M. C. & Kroemer, G. (2011). *Oncosuppressive functions of autophagy*, *Antioxidants and Redox Signaling* **14**, 2251-2269.

Mortimore, G. E. & Poso, A. R. (1987). *Intracellular Protein Catabolism and its Control During Nutrient Deprivation and Supply*, *Annual Review of Nutrition* **7**, 539-568.

Mortimore, G. E. & Schworer, C. M. (1977). *Induction of autophagy by amino-acid deprivation in perfused rat liver*, *Nature* **270**, 174-176.

Nakatogawa, H., Ichimura, Y. & Ohsumi, Y. (2007). *Atg8, a Ubiquitin-like Protein Required for Autophagosome Formation, Mediates Membrane Tethering and Hemifusion*, *Cell* **130**, 165-178.

Noda, N. N., Kumeta, H., Nakatogawa, H., Satoo, K., Adachi, W., Ishii, J., Fujioka, Y., Ohsumi, Y. & Inagaki, F. (2008). *Structural basis of target recognition by Atg8/LC3 during selective autophagy*, *Genes to Cells* **13**, 1211-1218.

Novak, I., Kirkin, V., McEwan, D. G., Zhang, J., Wild, P., Rozenknop, A., Rogov, V., Löhr, F., Popovic, D., Occhipinti, A., Reichert, A. S., Terzic, J., Dötsch, V., Ney, P. A. & Dikic, I. (2010). *Nix is a selective autophagy receptor for mitochondrial clearance*, *EMBO Reports* **11**, 45-51.

Pankiv, S., Clausen, T. H., Lamark, T., Brech, A., Bruun, J. A., Outzen, H., Øvervatn, A., Bjørkøy, G. & Johansen, T. (2007). *p62/SQSTM1 binds directly to Atg8/LC3 to facilitate degradation of ubiquitinated protein aggregates by autophagy*[S]*, *Journal of Biological Chemistry* **282**, 24131-24145.

Paz, Y., Elazar, Z. & Fass, D. (2000). *Structure of GATE-16, membrane transport modulator and mammalian ortholog of autophagocytosis factor Aut7p*, *Journal of Biological Chemistry* **275**, 25445-25450.

Pickford, F., Masliah, E., Britschgi, M., Lucin, K., Narasimhan, R., Jaeger, P. A., Small, S., Spencer, B., Rockenstein, E., Levine, B. & Wyss-Coray, T. (2008). *The autophagy-related protein beclin 1 shows reduced expression in early Alzheimer disease and regulates amyloid β accumulation in mice*, *Journal of Clinical Investigation* **118**, 2190-2199.

Pimkina, J., Humbey, O., Zilfou, J. T., Jarnik, M. & Murphy, M. E. (2009). *ARF induces autophagy by virtue of interaction with Bcl-xl*, *Journal of Biological Chemistry* **284**, 2803-2810.

Qu, X., Yu, J., Bhagat, G., Furuya, N., Hibshoosh, H., Troxel, A., Rosen, J., Eskelinen, E. L., Mizushima, N., Ohsumi, Y., Cattoretti, G. & Levine, B. (2003). *Promotion of tumorigenesis by heterozygous disruption of the beclin 1 autophagy gene*, *Journal of Clinical Investigation* **112**, 1809-1820.

Ravikumar, B., Duden, R. & Rubinsztein, D. C. (2002). *Aggregate-prone proteins with polyglutamine and polyalanine expansions are degraded by autophagy*, *Human Molecular Genetics* **11**, 1107-1117.

Ravikumar, B., Moreau, K., Jahreiss, L., Puri, C. & Rubinsztein, D. C. (2010). *Plasma membrane contributes to the formation of pre-autophagosomal structures*, *Nature Cell Biology* **12**, 747-757.

Rogov, V. V., Suzuki, H., Fiskin, E., Wild, P., Kniss, A., Rozenknop, A., Kato, R., Kawasaki, M., McEwan, D. G., Loehr, F., Guentert, P., Dikic, I., Wakatsuki, S. & Doetsch, V. (2013). *Structural basis for phosphorylation-triggered autophagic clearance of Salmonella*, *Biochem. J.* **454**, 459-466.

Rozenknop, A., Rogov, V. V., Rogova, N. Y., Löhr, F., Güntert, P., Dikic, I. & Dötsch, V. (2011). *Characterization of the interaction of GABARAPL-1 with the LIR motif of NBR1*, *Journal of Molecular Biology* **410**, 477-487.

Rubinsztein, D. C. (2006). *The roles of intracellular protein-degradation pathways in neurodegeneration*, *Nature* **443**, 780-786.

Rubinsztein, D. C., Codogno, P. & Levine, B. (2012). *Autophagy modulation as a potential therapeutic target for diverse diseases*, *Nature Reviews Drug Discovery* **11**, 709-730.

Rubinsztein, D. C., Mariño, G. & Kroemer, G. (2011). *Autophagy and aging*, *Cell* **146**, 682-695.

Scherz-Shouval, R., Shvets, E., Fass, E., Shorer, H., Gil, L. & Elazar, Z. (2007). *Reactive oxygen species are essential for autophagy and specifically regulate the activity of Atg4*, *EMBO Journal* **26**, 1749-1760.

Schmid, D., Pypaert, M. & Münz, C. (2007). *Antigen-Loading Compartments for Major Histocompatibility Complex Class II Molecules Continuously Receive Input from Autophagosomes*, *Immunity* **26**, 79-92.

Schmid, N., Eichenberger, A. P., Choutko, A., Riniker, S., Winger, M., Mark, A. E. & van Gunsteren, W. F. (2011). *Definition and testing of the GROMOS force-field versions 54A7 and 54B7*, *European Biophysics Journal* **40**, 843-856.

Shibata, M., Lu, T., Furuya, T., Degterev, A., Mizushima, N., Yoshimori, T., MacDonald, M., Yankner, B. & Yuan, J. (2006). *Regulation of intracellular accumulation of mutant huntingtin by beclin 1*, *Journal of Biological Chemistry* **281**, 14474-14485.

Shpilka, T., Weidberg, H., Pietrokovski, S. & Elazar, Z. (2011). *Atg8: an autophagy-related ubiquitin-like protein family*, *Genome Biol.* **12**, 226.

Shvets, E., Fass, E., Scherz-Shouval, R. & Elazar, Z. (2008). *The N-terminus and Phe52 residue of LC3 recruit p62/SQSTM1 into autophagosomes*, *Journal of Cell Science* **121**, 2685-2695.

Singh, R., Kaushik, S., Wang, Y., Xiang, Y., Novak, I., Komatsu, M., Tanaka, K., Cuervo, A. M. & Czaja, M. J. (2009). *Autophagy regulates lipid metabolism*, *Nature* **458**, 1131-1135.

Singh, R., Xiang, Y., Wang, Y., Baikati, K., Cuervo, A. M., Luu, Y. K., Tang, Y., Pessin, J. E., Schwartz, G. J. & Czaja, M. J. (2009). *Autophagy regulates adipose mass and differentiation in mice*, *Journal of Clinical Investigation* **119**, 3329-3339.

Sou, Y.-s., Tanida, I., Komatsu, M., Ueno, T. & Kominami, E. (2006). *Phosphatidylserine in Addition to Phosphatidylethanolamine Is an in Vitro Target of the Mammalian Atg8 Modifiers, LC3, GABARAP, and GATE-16*, *Journal of Biological Chemistry* **281**, 3017-3024.

Su, X., Wang, X., Liu, Q., Wang, P., Xu, C. & Leung, A. W. (2015). *The role of Beclin 1 in SDT-induced apoptosis and autophagy in human leukemia cells*, *International Journal of Radiation Biology* **91**, 472-479.

Sugawara, K., Suzuki, N. N., Fujioka, Y., Mizushima, N., Ohsumi, Y. & Inagaki, F. (2004). *The crystal structure of microtubule-associated protein light chain 3, a mammalian homologue of Saccharomyces cerevisiae Atg8*, *Genes to Cells* **9**, 611-618.

Suzuki, H., Tabata, K., Morita, E., Kawasaki, M., Kato, R., Dobson, R. C. J., Yoshimori, T. & Wakatsuki, S. (2014). *Structural basis of the autophagy-related LC3/Atg13 LIR complex: Recognition and interaction mechanism*, *Structure* **22**, 47-58.

Szeto, J., Kaniuk, N. A., Canadien, V., Nisman, R., Mizushima, N., Yoshimori, T., Bazett-Jones, D. P. & Brumell, J. H. (2006). *ALIS are stress-induced protein storage compartments for substrates of the proteasome and autophagy*, *Autophagy* **2**, 189-199.

Takamura, A., Komatsu, M., Hara, T., Sakamoto, A., Kishi, C., Waguri, S., Eishi, Y., Hino, O., Tanaka, K. & Mizushima, N. (2011). *Autophagy-deficient mice develop multiple liver tumors*, *Genes and Development* **25**, 795-800.

Tanida, I., Ueno, T. & Kominami, E. (2004). *Human light chain 3/MAP1LC3B Is cleaved at its carboxyl-terminal Met 121 to expose Gly120 for lipidation and targeting to autophagosomal membranes*, *Journal of Biological Chemistry* **279**, 47704-47710.

Thielmann, Y., Weiergräber, O. H., Mohrlüder, J. & Willbold, D. (2009). *Structural framework of the GABARAP-calreticulin interface - Implications for substrate binding to endoplasmic reticulum chaperones*, *FEBS Journal* **276**, 1140-1152.

Thurston, T. L., Ryzhakov, G., Bloor, S., von Muhlinen, N. & Randow, F. (2009). *The TBK1 adaptor and autophagy receptor NDP52 restricts the proliferation of ubiquitin-coated bacteria*, *Nature immunology* **10**, 1215-1221.

Tsukada, M. & Ohsumi, Y. (1993). *Isolation and characterization of autophagy-defective mutants of Saccharomyces cerevisiae*, *FEBS Letters* **333**, 169-174.

Vagin, A. & Teplyakov, A. (1997). *MOLREP: An Automated Program for Molecular Replacement*, *Journal of Applied Crystallography* **30**, 1022-1025.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. & Berendsen, H. J. C. (2005). *GROMACS: Fast, flexible, and free*, *Journal of Computational Chemistry* **26**, 1701-1718.

Van Wesenbeeck, L., Odgren, P. R., Coxon, F. P., Frattini, A., Moens, P., Perdu, B., MacKay, C. A., Van Hul, E., Timmermans, J.-P., Vanhoenacker, F., Jacobs, R., Peruzzi, B., Teti, A., Helfrich, M. H., Rogers, M. J., Villa, A. & Van Hul, W. (2007). *Involvement of PLEKHM1 in osteoclastic vesicular transport and osteopetrosis in incisors absent rats and humans*, *J. Clin. Invest.* **117**, 919-930.

Virgin, H. W. & Levine, B. (2009). *Autophagy genes in immunity*, *Nature Immunology* **10**, 461-470.

von Muhlinen, N., Akutsu, M., Ravenhill, B. J., Foeglein, A., Bloor, S., Rutherford, T. J., Freund, S. M. V., Komander, D. & Randow, F. (2012a). *LC3C, Bound Selectively by a Noncanonical LIR Motif in NDP52, Is Required for Antibacterial Autophagy*, *Mol. Cell* **48**, 329-342.

von Muhlinen, N., Akutsu, M., Ravenhill, B. J., Foeglein, Á., Bloor, S., Rutherford, T. J., Freund, S. M. V., Komander, D. & Randow, F. (2012b). *LC3C, Bound Selectively by a Noncanonical LIR Motif in NDP52, Is Required for Antibacterial Autophagy*, *Molecular Cell* **48**, 329-342.

Webb, J. L., Ravikumar, B., Atkins, J., Skepper, J. N. & Rubinsztein, D. C. (2003). *α-synuclein Is Degraded by Both Autophagy and the Proteasome*, *Journal of Biological Chemistry* **278**, 25009-25013.

Weidberg, H., Shpilka, T., Shvets, E., Abada, A., Shimron, F. & Elazar, Z. (2011). *LC3 and GATE-16 N Termini Mediate Membrane Fusion Processes Required for Autophagosome Biogenesis*, *Developmental Cell* **20**, 444-454.

Weidberg, H., Shvets, E., Shpilka, T., Shimron, F., Shinder, V. & Elazar, Z. (2010). *LC3 and GATE-16/GABARAP subfamilies are both essential yet act differently in autophagosome biogenesis*, *EMBO Journal* **29**, 1792-1802.

Wen, J., Arakawa, T. & Philo, J. S. (1996). *Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions*, *Analytical Biochemistry* **240**, 155-166.

Williams, A., Sarkar, S., Cuddon, P., Ttofi, E. K., Saiki, S., Siddiqi, F. H., Jahreiss, L., Fleming, A., Pask, D., Goldsmith, P., O'Kane, C. J., Floto, R. A. & Rubinsztein, D. C. (2008). *Novel targets for Huntington's disease in an mTOR-independent autophagy pathway*, *Nature Chemical Biology* **4**, 295-305.

Xie, Z., Nair, U. & Klionsky, D. J. (2008). *Atg8 Controls Phagophore Expansion during Autophagosome Formation*, *Molecular Biology of the Cell* **19**, 3290-3298.

Yang, L., Li, P., Fu, S., Calay, E. S. & Hotamisligil, G. S. (2010). *Defective hepatic autophagy in obesity promotes ER stress and causes insulin resistance*, *Cell Metabolism* **11**, 467-478.

Yang, Z. J., Chee, C. E., Huang, S. & Sinicrope, F. A. (2011). *The role of autophagy in cancer: Therapeutic implications*, *Molecular Cancer Therapeutics* **10**, 1533-1541.

Young, A. R. J., Chan, E. Y. W., Hu, X. W., Köchl, R., Crawshaw, S. G., High, S., Hailey, D. W., Lippincott-Schwartz, J. & Tooze, S. A. (2006). *Starvation and ULK1-dependent cycling of mammalian Atg9 between the TGN and endosomes*, *Journal of Cell Science* **119**, 3888-3900.

Yue, Z., Jin, S., Yang, C., Levine, A. J. & Heintz, N. (2003). *Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor*, *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15077-15082.

Zhang, Y., Goldman, S., Baerga, R., Zhao, Y., Komatsu, M. & Jin, S. (2009). *Adipose-specific deletion of autophagy-related gene 7 (atg7) in mice reveals a role in adipogenesis*, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 19860-19865.

Zhao, H., Laitala-Leinonen, T., Parikka, V. & Vaananen, H. K. (2001). *Downregulation of small GTPase Rab7 impairs osteoclast polarization and bone resorption*, *J. Biol. Chem.* **276**, 39295-39302.

Zheng, S., Clabough, E. B. D., Sarkar, S., Futter, M., Rubinsztein, D. C. & Zeitlin, S. O. (2010). *Deletion of the huntingtin polyglutamine stretch enhances neuronal autophagy and longevity in mice*, *PLoS Genetics* **6**.